Switching (or not) health seeking behavior: Evidence from rural Tanzania*

Lucia Corno†
Bocconi University

August 2008

Abstract

The aim of this paper is to understand the functioning of individuals' health seeking behavior. It studies empirically how and if economic agents adjust their health seeking behavior over time, after they have gained experience about the quality of the previous caregiver consulted. I find that agents seek medical care repeatedly from the same type of health provider, even if the treatments are ineffective. Specifically, they do not switch to formal health sector, even though the informal one has failed to treat their illness in previous years. The paper also investigates the determinants of illnesses, showing the relevant role of education in avoiding diseases. I apply the Wooldridge's (1995) procedure to deal with the sample selection bias problem, by exploiting past natural disasters as exclusion restriction. These effects are tested using a 4 years panel data from a household survey in Tanzania.

*I wish to thank Eliana La Ferrara for her guidance and support; Martina Bjorkman, Domenico Depalo, Andrea Gamba and the participants at the CSAE Conference in Oxford and at the EUDN Conference at Paris School of Economics for their very useful comments. I am also indebted to Kathleen Beegle for suggestions on the datasets used and to Paola Porta for sharing information on Kagera's diseases. All errors are my own.

†Correspondence: Bocconi University, IEP, via Roentgen 5, 20136, Milano, Italy. Email: lucia.corno@unibocconi.it

Keywords: health, Tanzania, formal, informal, switching, disaster
JEL classification: O12, I11, N37

1 Introduction

Illness is one of the most serious problems that can affect a household in African communities, where a considerable part of the population seeks medical care in the informal health sector, e.g. traditional healers and witch-doctors, or it doesn't receive any health treatments. Economics research (Björkman and Svensson 2006), international organizations (USAID 2006, WHO 2005, The World Bank 2004) and governments' efforts are mainly focused on the supply-side of the health market: how to improve health systems and how to increase the number and the quality of accessible medical services. In recent years, the growing attention in studying health care demand (Lindelow, 2002, Leonard 2007) has started to give some insights to better understand failures in the health systems of many developing countries. A limitation of this literature, however, is that it doesn't analyze the pattern of health seeking behavior over time. The main contribution of the present paper is to study individual's health seeking behaviors, conditional on the outcome of the previous therapy. It explores whether agents switch to a better quality health provider, after the failure in treating an illness on the part of the first caregiver consulted. The idea is to understand whether remaining sick after a treatment is a good incentive to switch to an alternative care. Understanding incentives/causes driving individuals towards more efficient health care is relevant to designing long term policy measures in the health systems. As main results, the paper shows that individuals, who sought care from informal caregivers don't switch to formal health cares even if they are still suffering from the same disease after the informal treatment. It emerges that the prior belief about what is the best medical care dominates over the illness status of individuals. The intuition under this result is straightforward: as long as the costs to test a new treatment are higher then the potential benefits coming from the new treatment, even a completely "rational" agent will choose not to experiment. This player gets stuck in a non-optimal equilibrium, simply because the cost of trying something else is higher compared with the potential benefits. In the paper, an agent has to compare the certain cost of formal health care (far away government-run clinics, long queues, etc.) with an uncertain benefit (the probability to heal). The confirmation of health seeking choices independently of their outcomes fosters the existence of unqualified doctors and health services. In addition, this work proposes a unique procedure to solve the selection bias problem, by exploiting past disasters as exclusion restriction, correlated with the probability of illness but not with the probability to look for formal health care.

The paper attempts to address three main questions. First, do agents switch to a formal caregiver, following only private information about their bad health status after the informal treatment? Second, what factors influence the demand for formal health care? Third, what are the main determinants of becoming ill?

Regarding the first question, I show that agents do not learn from their own past experience that it is better to switch to formal medical care after an ineffective informal therapy. Individuals choose once and for all at the beginning of the period and they do not update their believes, even if they have evidence of inefficient outcomes. The formal institutions I considered are hospitals, health centres, clinics and dispensaries and the informal ones are practitioners' homes, pharmacies, family homes and self-care. Ceteris paribus, to consult informal providers and to remain sick after the treatment decreases the probability to visit formal establishments in the next period by 11.5 percentage points.

A second result concerns the determinants involved in consulting formal or informal health care providers, conditional on reporting illnesses. The main determinant of the probability to seek formal care is the distance between the household and formal health establishments. As expected, more educated individuals tend to choose formal caregivers. Having access only to bad sources of drinking water, such as rain, lake and river water and living in an inadequate house, decreases the likelihood to seek formal medical care. Education, the quality of drinking water and living conditions also capture income and the ability to afford better services.

Finally, I show that disaster, such as drought, epidemic, insect and crop diseases, happened in a community six months prior to the survey increases the probability of illnesses. Years of education and bad living conditions are the main variables that determine one's possibility of becoming ill, affecting it respectively in a negative and in a positive way. Furthermore, women are more likely to report diseases compared to men.

The proposed analysis is tested using household panel data from Kagera, Tanzania. The Kagera Health Developing Survey has several features that make it particularly appropriate for studying individual's health seeking behavior. First, it contains detailed information on ill individuals as well as the type of illness reported, the decision to seek care, the type of health provider chosen and if respondents are still suffering from the same disease after a health visit. Second, more than one half of respondents reported an illness or an injury in the four weeks proceeding the survey. This high percentage provides a significant sample of ill individuals to analyze. Finally, the KHDS includes a community questionnaire. Hence, I exploit information on community shocks as an instrument to study the probability of illness and to solve the sample selection bias problem.

The paper is related to three main strands of the literature. First, it relies on studies on the determinants of health seeking behavior, such user fees (Gertler and Van Der Gaag 1990, Dasgupta and Gupta 2002), travel distance (Acton 1975), individual and household characteristics (Lindelow 2005). Case, Menendez and Ardington (2005) examine patterns of health seeking behavior prior to death among individuals in a South Africa district, finding that all adults who were ill prior to death sought treatment from a Western medical provider and, the fifty percent of them, from a traditional healer, suggesting that traditional medicine is seen as a complement to, rather than a substitute for, Western care. While previous works are mainly based on cross-section datasets and they adopt a static approach, this paper studies the health seeking behavior over time in a panel dataset, by conditioning the choice of health provider to the results of the previous therapy.

Second, the paper fits, for some aspects, into the more recent research highlights the role of social learning played in health seeking behavior. Das and Hammer (2004) evaluate how provider quality affects both the demand for health services and health outcomes, analyzing the competence of caregivers through vignettes and clinical observations in Delhi. Luke and Munshi (2007) assess the role of social affiliation, measured by caste in India, on household health care decisions and they propose a network-based explanation for why investments in health may differ across castes. Within this literature, the paper most related to my work is Leonard (2007). Using two separate datasets from Tanzania, he shows that households, by gathering information and by communicating with nearby households, learn about the best quality clinicians. Different from the latter, the idea of the present paper is to study the capacity to modify/switch behavior whenever new individual experience about provider's quality is acquired.[1] Psychological research indicates that some people have a cognitive bias that leads them to misinterpret new information as supporting previously held hypotheses.

Indeed, the results of the paper can be thirdly absorbed into the literature studies confirmation bias. Rabin and Schrag (1999) show that such confirmatory bias induces overconfidence: the agent may come to believe a false hypothesis despite receiving an infinite amount of information. The paper provide empirical evidence that agents confirm the choice of the first health provider consulted, no matter if past experience and new information have been acquired.

The reminder of the paper is organized as follows. Section 2 introduces some notions on the health system in Tanzania. Section 3 describes the data and illustrates the main pattern of health seeking decision. Section 4 shows the empirical strategy used and in section 5 I expose the main results. Section 6 concludes.

2 Health care system in Tanzania

Looking at other eastern African countries, Tanzania's health care system is relatively well-established. Health indicators are still slightly above the average for sub-Saharan Africa and the health sector is doing relatively well. Infant mortality rate fell from 99 deaths per 1,000 live births in 1999 to 68 per 1,000 in 2005. Children's malnutrition status also improved. Between 1999 and 2005, the incidence of stunting decreased from 44 percent to 38 percent, wasting from 5 to 3 percent and underweight from 29 to 22 percent. In contrast, maternal mortality remains high at 578 deaths per

[1] In the literature this is better known as the concept of "learning by doing" (Foster and Rosenzweig, 1995).

100,000 live births in 2005, life expectancy at births is 51 years old and according to the 2003-04 Tanzania HIV Indicator Survey (THIS) the 7% of adults where infected with HIV/AIDS. Malaria is the leading cause of death in Tanzania and the major public health concern, especially among pregnant women and children under five years (NBS and ORC Macro 2006).

The health sector in Tanzania improved after the independence from the United Kingdom in 1964, especially during the Arusha Declaration[2], when there was an emphasis for a more equal and efficient access to social services; before they were mainly basic and concentrated only in more developed urban centres. Free medical services were introduced and facilities were redirected towards rural and poorer areas. In 1991, the Government revised its approach towards private initiatives and it officially recognized the private activities of medical practitioners and dentists. Today, even if the government remains the main source of financing for the health sector, private organizations and qualified persons can actively participate in the development and management of health care services in Tanzania, with the approval from the Ministry of Health.

The health system has a pyramidal structure from services at district level to medical care at national level.[3] The lowest level of health care delivery in each district of the country is the village health service or health post. It essentially provides preventive services which can be offered in patients' homes, such as advice on personal habits and domestic hygiene (water, food), as well as the distribution of basic drugs. Usually, each health post has two workers chosen by the community who take short training courses before starting to provide services. The second stage of health services at district level is the dispensary. The dispensary provides care approximately for 6,000 to 10,000 people and supervises all the village health posts in its ward. Dispensaries in Tanzania are generally located in rural centres, serving population spread over a wide area. They are staffed by a rural medical aid (RMA) with one or two helpers. The RMA receives a 3 years course of training in anatomy,

[2] A declaration outlining Tanzania's policy on socialism and self-reliance, by Julius Nyerere.

[3] Ministry of Health and Social Welfare, The United Republic of Tanzania, National website, www.tanzania.go.tz., 2007.

physiology and hygiene with good grounding in diagnostic methods and treatment of common diseases, but they still offer limited types of surgery. Some dispensaries are managed by religious groups and, as a whole, they tend to be better equipped than those in the state sector, with higher standard of cleanliness and hygiene. The health centre provides a slightly higher level of medical care. It is expected to cater to 50,000 people, which is approximately the population of one administrative division. Most of health centres have a room for minor surgery and they provide 20-30 beds for inpatients including maternity cases. Each district is supposed to have a hospital. The district hospital is the base for staffing and supplying all rural units and it is to the district hospital that any difficult of serious case is referred. The regional hospital offers similar assistance like that agreed at the district level, but with specialists in various fields and additional services. Finally, at the top there are four referral hospitals which serve the whole country. Referral hospitals are managed by the Ministry of Health, while local governments are responsible for dispensaries, health centres and the district/regional hospitals. In this official classification clinics do not appear. They provide similar services to health centres and they are generally managed by NGOs or private organizations.

In addition to this formal structure of the health care delivery system, the sector includes individuals not recognized by the Ministry of Health such as practitioners working from their homes, traditional healers, quacks and witch-doctors. They serve as primary health care providers for a large part of the population. Witch-doctors have no formal training and acquired all of their practices from their parents who in turn learned from their grandparents. Their only diagnostic tools are patients' symptoms and for each ailment they prescribe a combination of songs, prayers, particular kinds of soil, and tree roots. In slums and rural area, sick poor people tend to consult them, since formal health establishments are far away and with long waiting lists; while informal caregivers generally visit people at home and so they are more accessible or easier to reach by all patients.

3 Data and Descriptive statistics

The data used for this study come from a research project conducted by the World Bank and the University of Dar es Salaam in the Kagera region, located in the North-western corner of Tanzania and bordering Uganda to the North and Rwanda and Burundi to the West. The prime objective of the survey was to measure the economic impact of fatal illness in the region and to propose cost-effective strategies to help survivors. Therefore, it includes a large set of questions on health care and health behaviors. The Kagera Health and Development Survey (KHDS) surveyed 816 households four times from 1991-1994, with an average interval between surveys of 6-7 months. Households were selected in the six districts of Kagera region (Karagwe, Bukoba Urban, Bukoba Rural, Muleba, Biharamulo and Ngara), mainly in rural villages (Figure 1).[4]

[Insert table 1]

The sample includes 19,009 individuals interviewed up to four times in 51 clusters. The analysis focuses on two dependent variables: the probability to seek formal health care in the four weeks prior the survey and the probability of illness. Table 1 provides information on the percentage of ill individuals by gender and age. In the sample, more than the half of individuals reported being ill or injured in the four weeks proceeding the survey. As expected, illness reporting was particularly high for the elderly (46 years or older) and for infants between 0-2 years old. The lowest percentage of ill individuals belonged to the class of children between 6-15 years old. Women were slightly more prone to report illness than men in all age categories, except for male children under 2 years old.[5]

[Insert table 2]

[4] In 2004 another round data was collected (KHDS 2004) but the paper presents evidence only on the first four years due to the lack of a key variable for my identification strategy in the last wave. The KHDS 2004 doesn't include the question on the individual health outcome (still sick or not) after a treatment.

[5] In what follows, I consider illness and injury as the same variable.

Evaluating patients' choice towards a specific caregiver requires data on the disease suffered as it allows to account for the tendency to seek formal care only for more serious illnesses. Given the impossibility of controlling directly for the disease for lack of data, I exploit all the information I have regarding type of symptoms. First, I grouped the reported symptoms in two broader categories, depending on how serious they are: "routine" and "not routine" symptoms (Luke and Munshi, 2007). The "routine" ones are the most common and most widespread such as acute diarrhea, chronic diarrhea, weight loss, weakness, vomiting, acute fever, chills, cough, sore throat, severe headache, skin rush and wound. The "not routine" are recurring fever, fainting, difficult breathing, productive cough, coughing blood, abdominal pain, pain on passing urine, genital sores, burn, fracture, child birth and mental disorder.[6] Table 2 describes the symptoms suffered by the sick agents in the sample. Almost 90% of individuals reported less severe symptoms. Among those, the prevalent is acute fever (23.55%), one of the main symptoms of malaria. Roughly 13% of the sample reports cough and about 10.89% is suffering from severe headache, followed by chills, wound and acute diarrhea. Approximately 21% of individuals reported to have other types of symptoms, different from the listed ones. Finally, among the not routine symptoms, the share of individuals having abdominal pain is around 6.5%. Second, I divided the symptoms reported in two other categories: cyclic and not cyclic, depending on the fact that they appear or not in regular intervals of time.[7] This distinction is relevant to understand the efficacy of a treatment. Cyclic diseases show up independently from the medical care chosen. In this case, the health status (sick or not sick) doesn't depend on the therapy but only on the illness's course. In the empirical part of the paper, I restrict the analysis to the sub-samples of routine, not routine and not cyclic symptoms.

[Insert table 3]

[6] The classification between routine and not routine symptoms has been costructed by an anonymous doctor working in Kagera region.

[7] The symptoms reported as not cyclic are: Acute Diarrhea, Weight Loss,Acute Fever, Skin Rash, Weakness, Chills, Productive Cough, Blood Cough, Urine Pain, Genitals, Sore Throat, Burn, Fracture, Wound, Child Birth (Harrison 2006).

Third, I infer a patient's illness by combining the first two symptoms reported by each sick respondent with the most spread illnesses in Kagera region. This additional check is relevant since for example one or more routine symptoms could even generate a serious disease. The first row of table 3 shows the most widespread diseases in Kagera, while the first column describes all the possible symptoms listed in the survey. Each illness is linked to the corresponding symptoms. In the table, 1 and 2 are, respectively, the first and the second symptom for that specific disease. The X represent all additional symptoms that could appear if the agent is suffering from the disease reported in the first row of the table. I infer an agent's predicted illness by considering the first two symptoms suffered by individuals. For example, if an individual reported acute fever as first symptom and chills as the second one, he might probably suffer from malaria.[8] In the empirical section, the estimates include predicted diseases as additional controls.

[Insert table 4]

The survey proceeds as follows: individuals who reported an illness in the four weeks prior to the survey were asked if anybody had been consulted for treating this illness and, consequently, which was the first place where the patient sought care. The treatment choices available in Kagera region are hospitals, health centres, dispensaries, clinics, pharmacies, practitioner's homes, and patient's homes. As table 4 (panel A) shows, the greater part of the sample (68.3%) doesn't consult anyone for their disease and this percentage is slightly higher for women. In slums and rural area, sick poor people don't seek health care because clinics are too far away and the queues too long. In many rural areas there are no government-run clinics or dispensaries, so the only solutions are often to give up a treatment or to opt for a quack. Splitting the sample by age groups, it emerges that the highest percentage of individuals receiving medical care are infants (40%). To analyze individual health seeking behavior, I aggregate the health establishment choices in two broader categories: informal and formal health care. Hospitals, health centres, clinics and dispensaries

[8] Table 3 has been costructed with the help of two anonymous doctors, one of whom, work in Kagera region.

are considered formal cares.[9] Informal health establishments include pharmacies, practitioner's homes, their own home and "other".[10] Table 4, panel B reports the percentages of sick agents who decide to consult someone for treating their illness: more than 13% of these choose informal health care. The reaction to illnesses could vary with individual characteristics, such as gender and age. More men than women tend to seek informal care and the highest percent of individuals who report an informal consultation is represented by elderly over 45 years. Among formal health care, roughly the 23% of infants have been visited in hospitals compared to the 30% of adults between 31 and 45 years old. Dispensaries are the most widespread formal establishments where individuals decided to seek treatments, conditional on illness and consultation, and they are particularly widespread among kids aged 3-5 years old. Approximately 10% of agents sought care in a health centre, while clinics are less attended. In the informal sector, individuals are mainly treated in traditional practitioner's homes and this percentage increases for the elderly. The survey also investigates if sick people look for health care for the same illness in any other establishments after the first visit. The 88% of individuals don't consult any additional caregivers for the same disease. This is a crucial information since to empirically study the respondent's health seeking behavior, we have to be sure that the health status after a health care visit (formal or informal) is a consequence of that specific treatment and not of many treatments within the same month.

[Insert figure 2]

The crucial point of the paper is to investigate health behaviors in the wave after a formal or informal health treatment. The formal or informal therapy didn't work whether individuals sick in the four week proceeding the survey are still ill from the same disease during the survey after a medical consultation. Figure 2 describes the timing of the questions asked in the survey. The questions are repeated for four waves.

[9] The formal health centres are all the institutions officially recognised by the State in Tanzania (Tanzania National Website, 2008: www.tanzania.go.tz/health.htlm).

[10] Partages, shops or private laboratories.

[Insert table 5]

Table 5 is a transition matrix showing agent's behavior from wave t-1 to wave t for those who were still suffering from the same disease during the survey in t-1. Only the 20 percent of respondents, who have visited an informal caregivers (or haven't received any care) and have remained sick afterwards, have switched towards formal health providers in the following wave. This shows that the 80 percent of the still sick informal seekers have confirmed their health provider's choice. While the 58 percent of the "formal care visitors" in t-1, for which the treatment didn't work, has shifted towards informal cares (or has opted for no treatment) in wave t. From table 5, I construct the main independent variable in the empirical section: INFORMAL/NO CARE DIDN'T WORK is a dummy taking value 1 if the respondent has consulted an informal providers in t-1 and he is still sick during the survey in t-1 (1013 observations) and equal 0 if he remained sick in t-1 although he visited a formal caregivers (384 observations).

[Insert table 6]

Table 6 shows a cost-benefit analysis as a tool to evaluate advantages and disadvantages of the formal health sector in Kagera. The table provides descriptive evidence that searching care in formal institutions increases the probability to heal after the treatment. The 47 percent of patients, who received care for the first time by an informal provider, is still suffering from the same disease compared to 40 percent of those remained sick after therapy in hospitals or clinics. The test for the equality of the two coefficients rejects the null hypothesis with a p-value of 0.01. The t-test is valid also considering the sub-group of routine symptoms, while the difference between formal and informal care is not statistically significant when considering individuals suffering from more serious diseases, suggesting that some chronic illnesses cannot be benefit neither from the informal, nor from the formal health sector. By comparing agents who seek formal caregivers with those who haven't consulted any health providers, sick individuals are 40% and 42% respectively, but the t-test is not statistically significant anymore. Considering only those suffering from not routine symptoms, I found a greater percentage of sick agents after a formal therapy compared to those who didn't opt for medical care. The latter result can be explained by a higher proportion of death among those affected by a serious disease who had not sought health care.

A second slot of results describes the fraction of ill individuals treated by a doctor by type of health care. The percent of patients who have been visited by a doctor is 19 percent for those seeking formal care and 0.4 percentage for the ones consulting informal health providers. So, why should rational agents look for medical care in the informal health sector? The last row of table 6 shows the main cost associated with formal health care: the reported distance between the household and the medical establishment individuals have consulted. The distance from the household and the informal health facilities visited by each member is 2 km, much lower than the distance necessary to reach formal establishments. In the empirical section, I compute the average distance, among household members, between the household and the formal/informal caregivers they visited. On average, the distance between a household and a formal establishment is equal to 5.6 percent km, while the distance to visit an informal practitioner is lower than 1 kilometer.

Table A1 in the appendix shows summary statistics of the sample. As explanatory variables I used individual and household characteristics. The individual features include age, gender, years of education, religion and health status. The best measure of income in the KHDS is the value of physical assets held by the household (including the value of land, business, equipment, livestock, and dwelling). In the empirical part, I measure wealth as the log value of durable assets per household. Further, I control for the source of drinking water for the family and the living conditions. The measure of unsafe water includes water from rivers, lakes, wells without pump and rain. 80% of individuals had access only to unsafe drinking water and the 25% lived in a dwelling composed of walls of mud or bamboo, with soil as the main flooring and by roofs made of grass or mud. Finally, I exploit a variable from the community questionnaire as exclusion restriction in the empirical strategy: disasters that happened in each community in the past six months, such as flood, drought, epidemic and crop diseases.

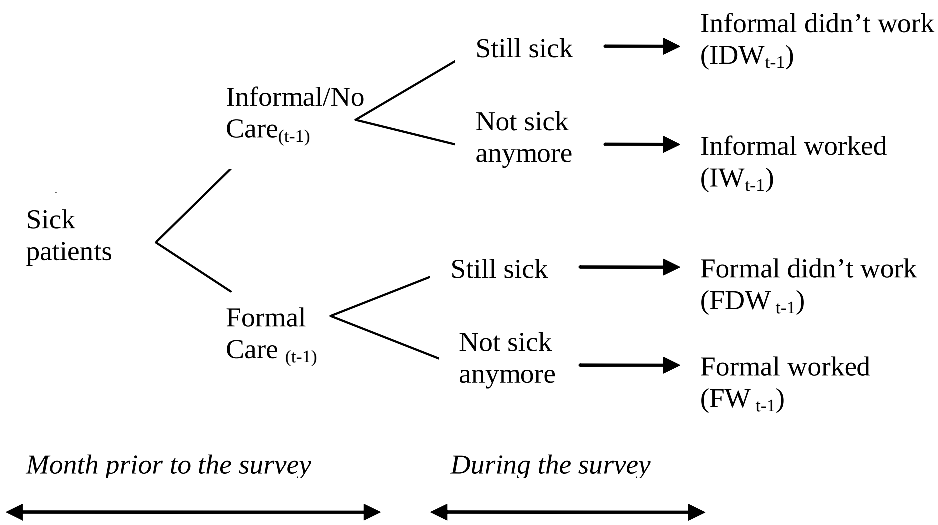
4 Empirical Strategy

4.1 Estimating equations

The main set of equations has the purpose to assess whether a patient's decision between formal and informal health care changes over time, depending on an agent's health status after an informal treatment. A patient would switch towards a formal health provider in the following wave when the informal one hasn't been effective in caring his/her disease. I adopt the following estimation strategy:

$$\begin{aligned} Y_{it} &= \alpha + \beta_1 IDW_{t-1} + \beta_2 IW_{t-1} + \beta_3 FW_{t-1} + \beta_4 Distance_{it} + \\ &\quad \beta_5 X_{it} + C_{it} + d_t + d_d + \varepsilon_{it} \end{aligned} \tag{1}$$

where Y_{it} is a binary variable equal one if the agent i seeks formal medical care and equal zero if he gets informal treatments or no treatments at time t. This variable is observed only if the individual is ill at time t, because, by construction, an healthy individual doesn't consult neither formal nor informal health providers. The selection bias problem involved in estimating equation 1 has been discussed and solved in paragraph 4.2. The KHDS investigates patient's health status after the formal/informal treatment, by asking whether he/she is still suffering from the same disease today (at the time of the interview). The following graph shows the four combinations of possible events and how the independent variables in equation 1 have been constructed.



The main variable of interest in equation 1 is INFORMAL/NO CARE DIDN'T WORK in t-1 (IWD_{t-1}), a dummy assuming value one if the agent i sought informal care or hadn't medical care in a month prior the survey in wave $t-1$, and he is still suffering from the same disease during the survey in the same wave. IW_{t-1} (Informal/no care worked) and FW_{t-1} (formal worked) are dummies equal 1 if agent visited an informal/formal provider and he is not sick anymore after the visit. In equation 1, the omitted category is represented by all the individuals who consulted formal doctors at time $t-1$ and are still sick during the survey at time $t-1$ (FDW_{t-1}). The paper aims to understand whether remaining sick after a treatment represents a good incentive to shift type of therapy. According to equation (1), the probability to look for formal treatment increases when $\beta_1 > 0$, suggesting that agents will switch from informal medical care to formal treatments, after a negative outcome (still sick) coming from the informal sector. In this case, a patient has consulted an informal provider in $t-1$, he has ascertained that it failed to treat his disease and therefore he updates his believes regarding the best quality caregiver. This updating will be translated in a different health provider's choice in the next wave. Agents confirm their health seeking choice whether $\beta_1 < 0$: they don't switch caregiver, even after the failure to treat diseases by the previous health provider consulted. X_{it} is a vector of individual and household characteristics, including age, sex, years of education, head of household, head's religion, household size and wealth, proxy as physical stock held by the family, living conditions (type of housing and access to safe water) and the dependency ratio (number of elderly and kids over household size). Notice that among the controls is the average distance between the household and the caregivers (formal or informal).[11] C_{it} is the children dummy equal 1 if the respondent is a kid younger than 12 years old, d_t and d_d are respectively time and district dummies and ε_{it} is the error term. Fixed effects at district level capture time-invariant features of the health market supply, such as the number of health facilities per district. Time t is the wave in which the survey has been done with an average interval between waves of 6-7 months.

The main concern with this empirical strategy is that the health status after the treatment could not necessarily be the result of success or failure of the previous treatment received. This could emerge for two reasons. First, patients might potentially visit many health care providers during their illnesses. The survey investigates the number of providers consulted since the beginning of the illness before checking respondent's health status at the time of the survey. Therefore, I estimate equation 1 also on the sample of those who only had received one medical advice for the same illness episode. Second, as explained in section 3, the health status could also depend on the type of symptoms and diseases suffered. In the empirical section, I restrict the estimates to non cyclic disease, routine and not routine symptoms and to predicted disease. Equation 1 will be estimated using a probit model.

4.2 Identification strategy

The main problem in identifying the probability of choosing a formal health provider is related to sample selection bias. The Kagera Health Development Survey reported information related to health care decisions only conditional to previous illnesses,

[11] The distance is self-reported. I calculate the average distance among household members who sought informal care and among those who sought formal care.

hence equation 1 can be estimated only for sick agents. There is a problem of selection whether some unobservable variables influence both the probability of illness and the utilization of health services (i.e. health endowment, health habits, etc.).

In cross-section datasets there have been attempts to solve this issue with a two-steps method such as the one proposed by Heckman (1979). The econometric literature does not yet provide a single solution to deal with sample selection bias in panel datasets. The first approach proposed is to control for unobserved time-invariant individual characteristics. Heckman and Macurdy (1980) argued that a fixed effect tobit model could be a solution for sample selection bias. Nijman and Verbeek (1992) consider a random effect model under the assumption of normality and serial independence of the idiosyncratic errors in both the selection and the main equation. Vella (1998) provides a review and additional references.

I control for sample selection bias following the procedure introduced by Wooldridge (1995, 2002). This method requires a standard probit or tobit model in the selection equation for each time period followed by a multivariate linear regression in the second stage equation. The Mills ratio coming from the first stage estimates are included as controls in the second stage linear estimate. I estimate the first stage equation with a probit model, for each time period t (four waves) as follows:

$$\Pr(ill)_{it} = 1[\beta_0 + \beta_1 Disaster_{ht} + \beta_2 X_{it} + \beta_3 \bar{X}_{it} + C_{it} + d_t + \varepsilon_{it} > 0] \qquad (2)$$

where $\Pr(ill)_{it}$ is the binary selection indicator equals to1 for agents reporting an illness or an injury in the month prior the survey for individual i at time t. $Disaster_{ht}$ is a dummy variable equal 1 if there was a shock in the community h in the six months prior to the survey, such as drought, epidemic, insect and crop diseases.[12] This variable is the exclusion restriction in the first stage equation correlated with the probability of illness, but not with the probability to seek formal

[12] The potential type of disasters mentioned in the questionnaire were: flood, drought, epidemic, insect, war and crop diseases. I didn't consider flood and war because of their potential endogeneity in supply side of health market (i.e. impact on the health facilities).

care (equation 1). Due to data limitations, we don't have an exclusion restriction at individual level: *Disaster* is constant within each cluster in the sample. Anyway, there are many reasons to think at it as a good exclusion restriction. In Tanzania, climatic conditions and unfavorable natural shocks, not only have a huge impact on agriculture (which accounts for almost half of GDP and employs 80% of the work force[13]) limiting cultivated crops, but they also influence people's health status, by worsening the already precarious living conditions. X_{it} represents individual and household controls as specified in equation 1, $\bar{X}_i$ is the average of each independent variables across time for the individual i[14],C_{it} is the children dummy equal 1 if the respondent is a kid younger than 12 years old, d_t represents four years time dummies and ε_{it} is the error term. The errors in the selection equation are assumed to be normally distributed, but to display arbitrary serial correlation and unconditional heteroskedasticity. I found reasonable not to include in $\bar{X}_i$ the average of the binary variables, because they present a good stability over time, and therefore, to decrease the risk of multicollinearity between the average and the single binary regressor.

The resulting estimates of the first stage equation are used to obtain the inverse Mills ratios, $\hat{\lambda}_{it}$ for all t and i.[15] Following Wooldridge (1995), I estimate the probability of seeking formal medical care Y_{it} by pooled OLS on X_{it}, $\bar{X}_i$, $\hat{\lambda}_{it}$ and $d_t{*}\hat{\lambda}_{it}$ for those observations for which $\Pr(ill)_{it} = 1$. Finally, the asymptotic variance of the estimated coefficients has been corrected by bootstrapping the standard errors.

To my knowledge this is the first attempt to solve the selection bias problem in panel data by empirically applying the Wooldridge's (1995) method.

[13] Central Intelligence Agency 2007.

[14] This is best viewed as the Mundlak approach (1978). A Chamberlain approach (1980) would replace $(X_{it}, \bar{X}_i)$ with X_i.

[15] The Mills ratio is $\hat{\lambda} = \phi(x\delta)/\Phi(x\delta)$, where $x\delta$ are the residuals of the equation 1.

5 Econometric results

5.1 Switching to formal health care

Table 7 reports estimated and marginal probit coefficients for a model where the dependent variable is equal 1 if the individual opts for formal establishments when sick.

[Insert table 7]

Columns 1 and 2 test the determinants of the probability to seek formal health care (hospitals, clinics, dispensaries, health centres), conditional on being ill in the four weeks prior to the survey. The distance between the household and the formal establishments shows a negative and significant coefficient: ceteris paribus one more kilometer of distance decreases the probability to seek formal care by 0.2 percentage points. On the contrary, the dependent variable increases with the number of years of education and the coefficient is significant at 5 percent level, suggesting the relevant role played by education in demanding higher standards of health care. Poorer households have a lower probability to be treated in hospitals or clinics: the dependent variable decreases with inadequate houses and with access only to unsafe water. The magnitude of the latter effect is substantial: conditional on illness, ceteris paribus, having access only to unsafe water (rain, river and lake) as the main source of drinking water reduces the likelihood to consult formal caregivers by 9 percentage points. As expected, household size decreases the dependent variable. Bad water, bad house and household size also capture a wealth effect, showing that, on average, wealthier households can afford better health services.

Columns 3 to 14 report probit estimates by conditioning the choice of a formal or informal medical care in time t to the outcome coming from the previous caregivers consulted. I test whether respondents switch health seeking behavior over time. The table reports only the coefficients of interest. In columns 3 and 4 I consider the full sample of individuals. Remember that the variable INFORMAL/NO CARE DIDN'T WORK$_{t-1}$ is equal 1 if the patient was ill four week prior to the survey, he consulted an informal health care provider (or he didn't receive a therapy) and he is still sick at the time of the interview. INFORMAL/NO CARE DIDN'T WORK$_{t-1}$ shows a negative and significant coefficient at 1 percent level. Looking at the magnitude of the coefficient we note that, ceteris paribus, having visited informal providers and being still sick reduces the probability to have formal care in the following year by 11.4 percentage points. Patients don't switch from informal to formal provider, even after the failure in treating their disease by the previous informal caregiver consulted. This result suggests lack of learning from own past experience in health seeking behavior. Moreover, the dependent variable decreases by 13.1 percentage points also when a patient has consulted an informal provider and he is not sick anymore (Informal worked $t-1$). These results confirm that remaining sick after a medical consultation isn't a good incentive to try alternative practitioners. The health status after a therapy doesn't count as a crucial factor in the health seeking choice in the next years. To isolate the effect of informal care from the one of not having care in t-1 without loosing observations, I include in all the estimates, a dummy variable equal 1 if the individual didn't consult anyone for his disease in t-1. The probability to seek formal care at time t decreases with the average distance between the household and the formal establishment (controlling also for the distance to informal practitioners) and the coefficient is statistically significant at 5 percent level. Compared to the previous specification, the head of household dummy is positive and significant, showing a higher probability to afford better services for the head of the family.

An individual's choice among different caregivers' options depends on the seriousness and on the type of symptom he is suffering from. Columns 5 to 10 report the findings by splitting the sample between routine (less serious), not routine (more serious) and not cyclic symptoms.[16] With all three restrictions the sign of the coefficient linked to INFORMAL/NO CARE DIDN'T WORK$_{t-1}$ is still negative and significant at 1% level, confirming the lack of learning. Columns 7 and 8 show an

[16]Note that in this case it is relevant to consider patients suffering from not cyclic disease in time t-1, to capture the type of symptom after the therapy. Cyclic symptoms appear from time to time and the agent's health status would not be the consequence of a treatment, but rather by the type of disease.

interesting result: for serious diseases, distance doesn't influence the probability to consult formal providers, while it is still statistically significant at 1 percent level for routine and not cyclic symptoms.

In columns 11 and 12 I include predicted diseases as additional controls. The dependent variable decreases with INFORMAL/NO CARE DIDN'T WORK$_{t-1}$ and patients suffering from malaria have a higher likelihood to look for formal health care, suggesting greater propensity to consult formal provider for serious disease. Holding other controls at the sample mean, the probability to get a formal care increases by 8.9 percentage points for agents at whom malaria has been inferred.

An explanation for the results so far could be driven by the fact that patients have consulted informal practitioners as first choice and then, they have sought other medical treatments within the same month. In this case, the concern would be that the outcome from the therapy is not the consequence of the first health care sought. To deal with this problem, I further restrict the sample to patients who only had one consultation for the same disease. Columns 13 and 14 show the estimates: once again β_1 remains negative and significant at 5 percent level.

5.2 Natural disasters and the probability of illness

To test the robustness of the results I control for sample selection bias, using as exclusion restriction a dummy variable equal 1 if a disaster happened in the community in the past six months, such as, drought, epidemic, insects and crop disease. By solving the selection bias problem, I demonstrate that the results so far are not driven by the selected sample of ill individuals.

[Insert table 8]

Columns 1 and 2 of table 8 report the first stage regression, using as a dependent variable the probability of illness in the past month. Individuals who live in a community hit by natural disasters have a higher probability to be sick in the following months. The coefficient is statistically significant at 1% level, suggesting that disaster could represent a good exclusion restriction in this context. Women have a higher probability of reporting illnesses in the four week proceeding the survey compared with men: being a woman increases the likelihood of illness by 4.4 percentage points. Education decreases the probability of illness with a coefficient statistically significant at 5 percent level. The head of household dummy has instead a positive effect on the likelihood to report an illness in the month prior to the survey. The head of a household is generally the one who has more opportunities to travel away from the family, increasing the risk of contracting diseases. As proxy of income I use an alternative wealth indicator: the amount of physical stocks held by the household.[17] Even controlling for district fixed effects, I show that the dependent variable decreases with the amount of physical assets held by the family, showing that richer individuals have a less probability to contract a disease. Finally, I investigate the relationship between diseases and household's characteristics. Extreme living conditions, such as a dwelling composed by wall of mud or bamboo, by earth as main floor and by roof made by grass or mud are associated with a higher probability of illness. To assess the magnitude of this coefficient, note that, ceteris paribus, having an inadequate house and drinking dirty water increase the probability of illness by 2.6 and 1.9 percentage points, respectively.

Columns 3 to 8 provide linear estimates on the probability to consult a formal caregivers controlling for sample selection bias and bootstrapping standard errors with 200 replications. With all the types of sample restrictions (routine, not routine, not cyclic symptoms and only one health visit) getting an informal health consultation and remaining sick has a negative and significant effect on the probability to seek formal health care in the next period. In the baseline specification, ceteris paribus the likelihood to switch health provider decreases by 11 percent with INFORMAL/NO CARE DIDN'T WORK$_{t-1}$. Patients don't update their believes on health seeking behavior after a negative outcome, once again revealing lack of learning. By controlling estimate for sample selection bias, I confirm the results reported in table 7, with a relevant role of distance, household size and access only to dirty water in the reducing the likelihood to seek formal health care.

[17] The physical stock considered are land, business, equipment, livestock, and dwelling.

As a robustness test for the pertinence of the exclusion restriction, I estimate equation 1 by including "disaster" as additional regressor. Table A2 in the appendix shows that a disaster happened in the community doesn't have any impact on the choice of formal health care.

6 Conclusion

The main contribution of the paper is to add a dynamic perspective in analyzing individual's health seeking behavior. Overall, the empirical evidence suggests that agents are biased towards one type of health care and they don't switch caregivers even if the treatment has failed to heal them. Patients behave without taking into account the private information on their health status. The paper also investigates how the choice between a formal health care provider (hospitals, health centres, dispensaries, clinics) and an informal one (pharmacy, practitioner's home, family homes, self-care or no-care) changes with individual and household characteristics. The main cost associated to the formal health sector is the distance between the household and the facility, while education positively affects the likelihood to have formal therapies. All the estimates are controlled for types of symptoms and diseases. By applying Wooldridge's procedure (1995) and by exploiting disasters which happened in the community as exclusion restriction (correlated with the probability of illness but not with the likelihood to have formal treatments), I attempt to solve the sample selection bias problem. The findings of the paper are not driven by the selected sample of sick individuals.

These results shed light on a relevant problem in Tanzania and they have important implications from a policy prospective. The individuals' health conditions are not only driven by the generally inefficient supply side of the health market, but even from the interesting structure of patient's demand. The success of a therapy is not an important factor involved in the choice of a specific caregiver. This feature fosters the existence of low quality and not qualified doctors and health services. The first step to enforce the demand for formal care is to promote education and to disseminate informative campaigns to overcome cultural bias towards informal caregivers. Second, a more capillary distribution of government-run health services is necessary. This is a very costly and long term solution, anyway. An alternative response could be the promotion of groups of official doctors in charge of visiting sick poor households in rural area from time to time.

In general this paper underlines the need for a better understanding of the demand for health services. Future research should explore more deeply the world of informal health institutions, by gathering data on estimated number of traditional healers and user fees.

7 References

Acton, J. P., *Nonmonetary Factors in the Demand for Medical Services: some empirical evidence*, Journal of Political Economy V. 83, n. 3, pp. 595-641, June 1975.

Alderman, H., Gertler, P., *Family Resources and Gender Differences in Human Capital Investments: The demand for Children's Medical Care in Pakistan*, Intrahousehold resource allocation in developing countries: Models, methods, and policy, pp. 231-48, Johns Hopkins University Press for the International Food Policy Research Institute, 1997.

Althaus, F. A., *Work in progress: The expansion of Access to Abortion services in South Africa following Legalization*, International Family Planning Perspectives, Vol.26, No. 2, pp.84-86, June 2000.

Banks, J. S., Sundaram, K. R., *Switching cost and the Gittins index,* Econometrica, Vol. 62, No. 3, pp. 687-694, May 1994.

Beegle K., De Weerdt J., Dercon S., *Kagera Health and Development Survey (KHDS), 2004, Basic Information Document*, mimeo, The World Bank, 2006b.

Björkman M., Svensson J. *Power to the People: Evidence from a Randomized Experiment of a Community Based Monitoring Project in Uganda*, CEPR Discussion Paper No.6344, 2007

Börgers T., *On the relevance of evolution and learning to economic theory*, Economic Journal 106, 1274-1385, 1996

Case A., Menendez A., Ardington C., *Health Seeking Behavior in Northern KwaZulu-Natal*, Princeton University working paper n.146, 2005.

Central Intelligence Agency (CIA), *The World Fact Book*, Tanzania, 2007, https://www.cia.gov/libr world-factbook/index.html

Conley, G. T., Udry C. R., *Learning About a New Technology: Pineapple in Ghana*, Yale Economic Growth Center Discussion Paper N. 817, July 2000.

Cretin, S., Keeler, E. B., *Uses of Cost -benefit Analysis: Editorial*, Journal of Health Economics v6, n3, pp. 275-278, September 1987.

Das, J., Hammer J., *Strained Mercy: The Quality of medical care in Delhi*, Policy Research Working Paper Series: 3228, The World Bank, 2004.

Das, J., Hammer, J., *Which Doctor? Combining Vignettes and Item-Response to Measure Doctor Quality*, Journal of Development Economics, 78, 348-383, 2005

Dasgupta, P., Gupta, I., *Demand for curative health care in rural India: Choosing between Private, Public and No care*, Working paper series N. 82, National Council of Applied Economic Research, January 2002.

Dupas, P. *The impact of Conditional In-Kind Subsidies on Preventive Health Behaviors: Evidence from Western Kenya*, EHESS-PSE, Paris, July 11, 2005.

Economic Development Initiative, *Kagera Rural Core Welfare Indicator Questionnaire (CWIQ), Baseline survey on poverty, Welfare and Services in Kagera Rural District,* April 2004.

Foster, A. D., Rosenzweig, M. R., *Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture*, Journal of Political Economy v103, n6 pp. 1176-1209, December 1995.

Fauci A.S, Braunwald E., Kasper D. L., Hauser L. S., Longo L. D., Jameson L., Loscalzo J. *Harrison's Principles of Internal Medicine*, 17th Edition, McGraw-Hill, June 2008.

Gertler, P., Van Der Gaag, J., *The willingness to Pay for Medical Care*, John Hopkins University Press for the World Bank, 1990.

Gittins, J. C., Jones, M. D., *A dynamic allocation index for the discounted multiarmed bandit problem*. Biometrika 66, pp. 561-565. 430, 1976.

Green, M., *Public reform and the privatization of poverty: some institutional determinants of health seeking behavior in southern Tanzania*, Culture, Medicine and Psychiatry 24, pp. 403-430, 2000.

Greene, W., *Estimating Econometric Models with Fixed Effects*, New York University, Leonard N. Stern School of Business, 2001.

Heckman, J., *Sample Selection Bias as a Specification Error*, Econometrica Vol.4, No. 1., pp. 153-162, Jan., 1979.

Heckman, J., Macurdy, T. E., *A Life Cycle Model of Female Labour Supply*, The Review of Economic Studies, Vol. 47, No 1, Ecometrics Issue, pp. 47-74, Jan, 1980.

Kasper, D., Braunwald E., Hauser, S., Longo, D., Jameson, J. L., Fauci, A., Harrison's Principles of Internal Medicine, The McGraw Hill, 16th Edition, 2006.

Kyriazidou, E., *Estimation of a Panel Data Sample Selection Model*, Econometrica, Vol. 65, No 6, pp 1335-1364, November 1997.

Leonard K., *Learing in health care: evidence of learning about clinician quality in Tanzania,* Economic Development and Cultural Change, 55 (3) pp.531-555, 2007.

Leonard K., Mliga G. M., Mariam D. H, *Bypassing Health Centres in Tanzania: Revealed Preferences for Observables and Unobservable Quality*, Journal of African Economies v11, n4, pp. 441-71, December 2002.

Leonard K., Masatu M. C., Vialou, A., *Getting Doctors to do their best: Ability, Altruism and Incetives*, Journal of Human Resources Vol.42 (3) 2007, pp. 682-700, 2005.

Lindelow, M., *The Utilization of Curative Healthcare in Mozambique: Does Income Matter?*, Journal of African Economics, Volume 14, Number 3, pp. 432-482, 2005.

Luke, N., Munshi, K., *Social Affiliation and the Demand for Health Services: Caste and Child Health in South India*, Journal of Development Economics 83(2):256-279, 2007.

Ministry of Health and Social Welfare, The United Republic of Tanzania, National website, www.tanzania.go.tz., 2007.

Munshi, K., *Social Learning in a heterogenous population: technology diffusion in the Indian Green Revolution*, Journal of Development Economics 73, 185-213, 2004.

National Bureau of Statistics and Macro International, *Tanzania Service Provision Assessment Survey 2006*, Dar es Salaam, Tanzania 2006.

Nijman, T, Verbeek, M., *Testing for Selectivity Bias in Panel Data Models*, International Economic Review, Vol. 33, No. 3, pp. 681-703, August 1992.

Rabin M., Schrag J. L., *First Impressions Matter: A Model of Confirmatory Bias,* The Quarterly Journal of Economics, Vol. 114, No. 1, pp. 37-82, 1999

Sahn D. E., Youger, S. D., Genicot, G., *The Demand for Health Care Services in Rural Tanzania*, Oxford Bulletin of Economics and Statistics, 65, 2, 0305-9049, 2003.

Semykina A., Wooldridge J., *Estimating panel data models in the presence of endogeneity and selection: theory and application*, working paper, June 2005.

Tanzania Government and ORC Macro, “Tanzania HIV/AIDS Indicator Survey 2003-04”, Tanzania Commission for AIDS and National Bureau of Statistics, Dar es Salaam, Tanzania, 2005.

Tanzania Government and ORC Macro, “Tanzania Demographic and Health Survey 2004-2005”, National Bureau of Statistics, Dar es Salaam, Tanzania, 2005.

The World Bank, "User’s Guide to the Kagera Health and Development Survey Datasets", mimeo, 2004.

USAID, Report to Congress: *Health-Related research and development activities at USAID*, June 2006.

Vella, F., *Estimating Models with Sample Selection Bias: A survey*, The Journal of Human Resources XXXIII, 1998.

WHO and European Observatory on Health Systems and Policies Series, *Purchasing to improve health system performance*, Edited by Josep Figueras, Ray Robinson, Elke Jakubowski, Open University Press, 2005.

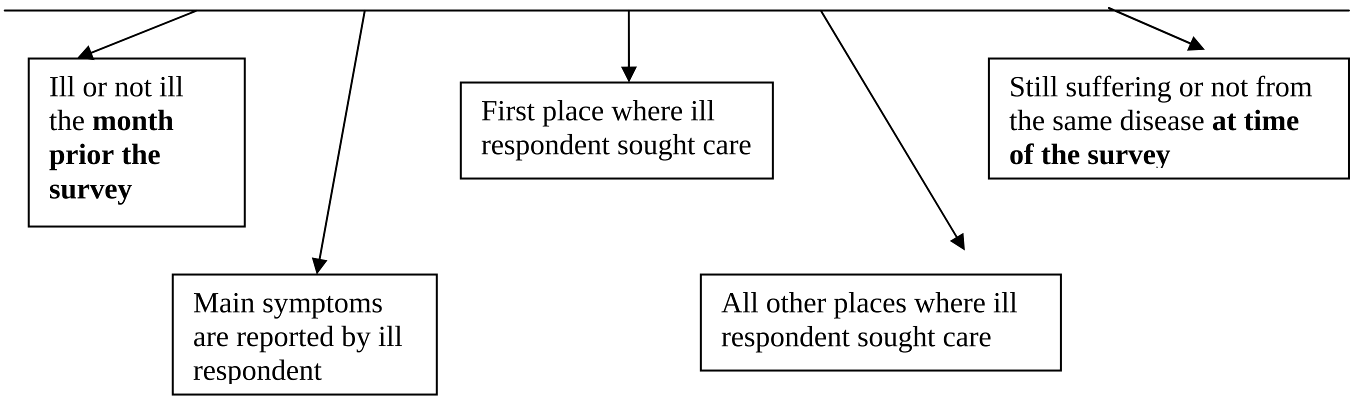
Wooldridge, J. M., *Selection corrections for panel data models under conditional mean independence assumptions*, Journal of Econometrics 68 pp.115-132, 1995.

Wooldridge, J. M., *Econometric analysis of cross section and panel data*, Massachusetts Institute of Technology, 2002.

Figure 1: Location of the KHDS clusters in Kagera Region, Tanzania

Source: The World Bank, 2004

Figure 2: Timing of the questions in the KHDS 1994-1997



Source: The World Bank, 2004

Tables

Table 1: Percent of individual reporting illness and injury in the four weeks prior to survey, by age and gender

Age	Total	Male	Female
0-2	63.63	64.02	63.17
3-5	48.48	47.96	48.99
6-15	40.59	40.52	40.67
16-30	45.87	42.87	48.61
31-45	55.73	51.27	58.92
>45	68.88	66.12	71.14
Total	51.24	49.77	52.60

Source: author's calculation on KHDS

Table 2: Symptoms of those reporting illness or injury in month prior to survey

Symptom	*%*	*Symptom*	*%*
Routine	88.21	**Not Routine**	11.79
Acute Diarrhea	3.7	Recurring fever	2.25
Chronic Diarrhea	0.11	Productive Cought	1.84
Weight Loss	0.05	Coughing Blood	0.08
Vomiting	0.85	Difficulty breathing	0.38
Acute Fever	23.55	Abdominal pain	6.51
Sever Headache	10.89	Pain on passing urine	0.02
Chills	8.21	Genital Sores	0.02
Cough	12.95	Burn	0.21
Sore throat	0.15	Fracture	0.19
Wound	4.06	Child birth	0.17
Weakness	1.68	Mental disorder	0.03
Skin Rash	1.22	Fainting	0.09
Other	20.79		

Source: author's calculation on KHDS.

Table 3: Symptoms reported associated with most widespread diseases in Kagera

Symptoms	Diseases													
	Malaria	Acute Respiratory infection	Helmintiasis	Skin Disease	AIDS	TB	Trauma	Veneral Disease	Child Birth	Diarroheal diseases	Mental illnesses	Urinary Tract Infection	Migraine	Burn
Acute Diarrhea			2			x				1				
Chronic Diarrhea					2									
Weight Loss					1	2	x					2		
Acute Fever	1	1				x								
Recurr. Fever					x		x	2						
Skin Rash				1	x									
Weakness	x					x	x						x	
Headache													1	
Fainting							x							
Chills	2					x								
Vomiting										2			2	
Cough		2												
Prod. Cough		x												
Coughing blood						1								
Urine Pain												1		
Genitals sores								1						
Mental Disorders	x										1			
Abdominal pain			1		x	x								
Sore Throat		x												
Breathing		x												
Burn														1
Fracture							2							
Wound							1							
Child Birth									1					

Note: 1=Main symptom in the disease, 2=second important symptom in the disease, X=additional symptoms eventually present in the disease.

Source: Fauci et al. 2008.

Table 4 : Choice of health care provider on sample of ill individuals

Place	*Total*	*By gender* Male	Female	*By age* 0-2	3-5	6-15	16-30	31-45	>45
Panel A									
Look for care	31.7	33.2	30.4	40.5	27.2	28.8	32.3	32.3	32
No Care	68.3	66.8	69.6	59.5	72.8	71.2	67.7	67.7	68.0
No.Obs	8770	4119	4651	991	828	2408	1951	988	1604
Panel B									
Informal care	*13.3*	*14.1*	*12.6*	*11.9*	*13.8*	*11.4*	*13.3*	*14.7*	*15.8*
Pharmacy	1.3	1.2	1.3	1	0.4	1.3	1.0	2.5	1.4
Practitioner	6.7	7.7	5.7	6.7	8.0	4.8	6.4	7.8	8.6
Own home	4.3	4.4	4.2	3.5	4.4	3.9	4.8	3.5	5.5
Other	1	0.7	1.3	0.7	0.9	1.4	1.3	0.9	0.4
No.Obs	370	178	192	48	31	79	84	47	81
Formal Care	*86.7*	*85.9*	*87.4*	*88.0*	*86.2*	*88.6*	*86.7*	*85.3*	*84.2*
Hospital	25.2	22.9	27.4	23.2	19.6	21.9	28.6	29.8	26.6
Health Centre	10.4	11.1	9.7	12.5	7.6	10.4	9.4	11.6	10.6
Dispensary	50	51.1	49	49.4	57.8	55.2	48.5	43.0	46.5
Clinic	1.1	0.8	1.3	3.0	1.3	1.2	0.2	0.9	0.6
No. Obs.	2410	1236	1174	353	194	615	545	272	431

Source: author's calculation on KHDS

Table 5: Switching between formal care and informal/no care from wave t-1 to wave t for those still ill after a therapy

Wave t	*Wave t-1* Informal/No Care		Formal Care		Total
	Obs.	%	Obs.	%	Obs
Informal/No Care	815	80	224	58	1039
Formal Care	198	20	160	42	358
Total	1013	100	384	100	1397

Source: author's calculation on KHDS

Table 6: Benefits and Costs of formal health care

Sample	*Formal* μ_F	*Informal* μ_I	*P-value* μ_F-μ_I=0	*Obs.*	*No care* μ_{NC}	*P-value* μ_F-μ_{NC}=0	*Obs.*
Fraction of ill individuals in the months prior the survey, who are still ill during the survey							
Pooled Sample	0.40	0.47	0.01	2781	0.42	0.13	8401
Routine Symp.	0.39	0.45	0.03	2349	0.42	0.01	7431
Not Routine Symp.	0.47	0.54	0.24	432	0.40	0.04	970
Fraction of ill individuals, treated by a doctor							
Pooled Sample	0.19	0.04	0.00	2661	-	-	-
Distance (Km) between the facility and the household [(a)]							
Pooled Sample	4.9	2	0.00	2782	-	-	-

Source: author's calculation on KHDS

[(a)] The distance from health facility to household is equal 0 if agents seek care in their own homes.

Table 7: Probability of care by formal providers if the informal sector didn't work

Dependent variable=1 if an individual seek formal care

	Baseline				*Type of symptoms*						*Type of diseases*		*Single visit (t-1)*	
					Routine		*Not routine*		*Not cyclic symp. (t-1)*					
	Coeff.	*Marg.* [(a)]	*Coeff.*	*Marg.*	*Coeff.*	*Marg.*	*Coeff.*	*Marg.*	*Coeff.*	*Marg.*	*Coeff.*	*Marg.*	*Coeff.*	*Marg.*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Infomal/no care didn't work $_{(t-1)}$			-0.352**	-0.115**	-0.306**	-0.099**	-0.877*	-0.297**	-0.334**	-0.108**	-0.338*	-0.121*	-0.470**	-0.165**
			(0.147)	(0.045)	(0.156)	(0.048)	(0.454)	(0.134)	(0.168)	(0.052)	(0.196)	(0.067)	(0.211)	(0.065)
Formal worked $_{(t-1)}$			-0.011	-0.004	-0.021	-0.007	0.070	0.026	-0.050	-0.017	-0.041	-0.015	0.006	0.002
			(0.085)	(0.029)	(0.091)	(0.030)	(0.267)	(0.100)	(0.099)	(0.033)	(0.110)	(0.040)	(0.094)	(0.036)
Informal/No Care worked $_{(t-1)}$			-0.400***	-0.132***	-0.375**	-0.122**	-0.743	-0.257*	-0.400**	-0.130**	-0.344*	-0.124*	-0.150	-0.056
			(0.147)	(0.046)	(0.155)	(0.048)	(0.473)	(0.149)	(0.168)	(0.052)	(0.195)	(0.068)	(0.199)	(0.073)
Distance to Formal	-0.007***	-0.002***	-0.019***	-0.006***	-0.019**	-0.007***	-0.006	-0.002	-0.023***	-0.008***	-0.014*	-0.005*	-0.018**	-0.007**
	(0.003)	(0.001)	(0.005)	(0.002)	(0.006)	(0.002)	(0.020)	(0.007)	(0.007)	(0.002)	(0.007)	(0.003)	(0.009)	(0.003)
Education	0.022**	0.008**	0.006	0.002	0.007	0.002	-0.004	-0.001	0.017	0.006	0.014	0.005	0.027	0.010
	(0.009)	(0.003)	(0.013)	(0.004)	(0.013)	(0.005)	(0.041)	(0.015)	(0.015)	(0.005)	(0.016)	(0.006)	(0.021)	(0.008)
HH Head	0.113	0.039	0.176*	0.062*	0.127	0.044	0.485*	0.186*	0.241**	0.084**	0.205*	0.077*	0.299*	0.117*
	(0.069)	(0.025)	(0.091)	(0.033)	(0.097)	(0.034)	(0.279)	(0.108)	(0.108)	(0.039)	(0.116)	(0.044)	(0.158)	(0.062)
Bad water	-0.252***	-0.090***	-0.271***	-0.096***	-0.252**	-0.088***	-0.517**	-0.199**	-0.256***	-0.090***	-0.361***	-0.138***	-0.399**	-0.156***
	(0.047)	(0.017)	(0.064)	(0.024)	(0.068)	(0.025)	(0.223)	(0.087)	(0.076)	(0.028)	(0.086)	(0.034)	(0.102)	(0.040)
Bad house	-0.095*	-0.032**	-0.102	-0.035	-0.092	-0.031	-0.169	-0.062	-0.135*	-0.045*	-0.126	-0.046	-0.098	-0.037
	(0.049)	(0.016)	(0.067)	(0.022)	(0.072)	(0.023)	(0.199)	(0.072)	(0.078)	(0.025)	(0.086)	(0.031)	(0.116)	(0.044)
HH size	-0.010*	-0.004*	-0.004	-0.001	-0.003	-0.001	-0.008	-0.003	-0.010	-0.004	0.005	0.002	-0.014	-0.005
	(0.005)	(0.002)	(0.007)	(0.003)	(0.008)	(0.003)	(0.023)	(0.009)	(0.009)	(0.003)	(0.010)	(0.004)	(0.013)	(0.005)
Predicted Malaria											0.230*	0.087*		
											(0.118)	(0.046)		
Observations	7384		2795		2477		318		1476		1570		955	
Wald chi2	277.87		164.85		153.1		40.13		121.7		119.2		59.54	
Prob > chi sq.	0.00		0.00		0.00		0.07		0.00		0.00		0.00	

Notes: * denotes significance at the 10 percent level; ** at the 5 percent level; *** at 1 percent level. Probit Estimates. Standard errors in parenthesis and corrected for heteroskedasticity and clustering of the residuals at the household level.

(a) Marginal probit coefficients are calculated at the mean. For dummies, marginal effect is calculated for discrete change from 0 to 1.

Estimates controlled for not having care in t-1, distance to infomal care, age, age sq., gender, head's religion, log of physical stock, household size, dependency ratio, wave dummies and district dummies.

.

Table 8: Probability of illness and of formal care, controlling for sample selection

Dependent variable=	1 if ill		1 if Formal Care					
			Baseline	*Type of symptom*			*Type of disease*	*Single visit $_{(t-1)}$*
	Coeff.	*Marg.* [(a)]		*Routine*	*Not Routine*	*Not cyclic $_{(t-1)}$*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Disaster [(b)]	0.072***	0.029***						
	(0.026)	(0.010)						
Informal/no care didn't work $_{(t-1)}$			-0.118**	-0.102**	-0.287*	-0.115**	-0.122*	-0.146**
			(0.048)	(0.051)	(0.157)	(0.056)	(0.070)	(0.069)
Formal worked $_{(t-1)}$			-0.011	-0.012	-0.028	-0.022	-0.027	-0.002
			(0.033)	(0.035)	(0.095)	(0.041)	(0.042)	(0.038)
Informal/No Care worked $_{(t-1)}$			-0.136***	-0.122**	-0.264	-0.136**	-0.129*	-0.049
			(0.047)	(0.052)	(0.176)	(0.058)	(0.070)	(0.073)
Distance to Formal			-0.004***	-0.004***	-0.001	-0.005***	-0.004**	-0.004*
			(0.001)	(0.001)	(0.006)	(0.001)	(0.002)	(0.002)
Female	0.111***	0.044***	-0.052*	-0.053*	0.025	-0.039	-0.075*	-0.041
	(0.026)	(0.010)	(0.027)	(0.032)	(0.098)	(0.037)	(0.045)	(0.056)
Education	-0.027**	-0.011**	0.019	0.019	0.005	0.042	0.041	0.024
	(0.012)	(0.005)	(0.018)	(0.019)	(0.071)	(0.022)	(0.026)	(0.037)
HH Head	0.175***	0.069***	0.082	0.083	-0.043	0.074	0.056	0.240
	(0.045)	(0.018)	(0.072)	(0.083)	(0.253)	(0.090)	(0.108)	(0.160)
HH Size	-0.003	-0.001	-0.122***	-0.140***	0.034	-0.010	-0.004	-0.023
	(0.010)	(0.004)	(0.037)	(0.041)	(0.050)	(0.017)	(0.023)	(0.030)
Bad water	0.028	0.011	-0.142***	-0.129***	-0.213**	-0.133***	-0.195***	-0.230***
	(0.031)	(0.012)	(0.028)	(0.031)	(0.084)	(0.034)	(0.038)	(0.049)
Bad house	0.055*	0.022*	-0.029	-0.021	-0.044	-0.048	-0.050	-0.003
	(0.030)	(0.012)	(0.029)	(0.029)	(0.104)	(0.031)	(0.038)	(0.063)
Predicted Malaria							0.103**	
							(0.047)	
Observations	18340		2687	2380	307	1943	1505	900
R sq.			0.0877	0.0883	0.3983	0.0938	0.15	0.174
Pseudo R sq.	0.0449							

Note: standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Columns 1 and 2 show probit estimates pooling the four waves together, altough to costruct the Mills ratios I separately estimate the probability of illness in each wave. Columns 3 to 8 report linear estimates.

Estimates controlled for not having care in t-1, distance to infomal care, age, age sq., log of physical stock, head's religion, dependency ratio, wave and district dummies

(a) Marginal probit coefficients are calculated at the mean. For dummies, marginal effect is calculated for discrete change from 0 to 1.

(b) Disaster includes drought, epidemic, insect and crop diseases.

Appendix Tables

Table A1: Definition of variables and descriptive statistics

Variable	Definition	Obs	Mean	S.D.
Informal/No care didn't work $_{(t-1)}$	Equal 1 if an ill individual in a month prior the survey consulted an informal provider at wave t-1 and he is still ill during the survey at wave t-1	2795	0.293	0.46
Formal didn't work $_{(t-1)}$	Equal 1 if an ill individual in a month prior the survey consulted an informal provider at wave t-1 and he is still ill during the survey at wave t-1	2795	0.211	0.41
Informal worked $_{(t-1)}$	Equal 1 if an ill individual in a month prior the survey consulted an informal provider at wave t-1 and he isn't ill during the survey at wave t-1	2795	0.359	0.48
No care $_{(t-1)}$	Equal 1 if an individual didn't have health care at wave t-1	2797	0.61	0.49
Formal$_{(t)}$	Equal 1 if formal provider (hospital, health centre dispensary clinic), 0 if informal (pharmacy, patient’s home, other, self care) at wave t	7384	0.324	0.47
Distance to formal care	Average distance (km) between the household and the formal health establishments visited by HH's members	7384	5.509	7.72
Distance to informal care	Average distance (km) between the household and the informal health establishments visited by HH's members	7384	0.12	0.7
Ill	Equal to 1 if individual reported ill or injury and 0 if not, in a month prior to survey	19009	0.512	0.5
Female	Equal 1 if female, 0 otherwise	19009	0.521	0.5
Age	Age in years	19009	22.47	20.2
Education	Years of education	19009	2.714	3.11
HH head education	Years of education of the HH head	19009	4.952	3.22
HH head	Equal 1 if head of household, 0 otherwise	19009	0.173	0.38
HH Catholic	Equal to 1 if the HH Head is catholic and 0 if he is muslim, protestant, other christian, traditional and other religions	19009	0.594	0.49
HH size	Number of member in each hh	19009	8.091	4.11
Log Phys. Stock	Log value of physical asset	19009	13.1	1.59
Bad water	Equal 1 if the source of drinking water for household river/lake, well without pump, rain 0 otherwise	19009	0.8	0.4
Bad house	Equal 1 if inadequate house (bad wall bad floor bad roof), 0 otherwise	19009	0.253	0.43
Dependecy Ratio	No. of elderry and kids in the household over HH size	19009	0.479	0.2
Children	Equal 1 if children younger than 12 years old	19009	0.396	0.49
Disaster	Equal 1 if there was a disaster in the past 6 months in the community, such as: drought, epidemic, insect, crop diseases.	18340	0.56	0.5

Source: author's calculation on KHDS

Table A2: Impact of disaster on the probability to look for formal care

Dependent variable=1 if an individual seek formal care	
Disaster [b]	0.041
	(0.076)
Informal/no care didn't work $_{(t-1)}$	-0.118
	(0.048)**
Formal worked $_{(t-1)}$	-0.010
	(0.032)
Informal/No Care worked $_{(t-1)}$	-0.135
	(0.047)***
Distance to Formal	-0.004
	(0.001)***
Observations	1159
R squared	0.0877

Note: standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Linear Estimates.
As additional controls I included age, age sq., gender, educatioon, distance to informal care household head, hh size, dep. ratio, wave and district dummies plus the average of all covariates, except for the dummies (Wooldridge 1995).

(b) Disaster includes drought, epidemic, insect and crop diseases.