Subsidies, Information, and the Timing of Children's Health Care in Mali

Anja Sautmann*, Samuel Brown† and Mark Dean‡

March 2016

Abstract

We study the impact of subsidies (which remove cost barriers) and healthworker visits (which remove informational barriers) on over- and underuse of primary care, using a randomized control trial across 1532 children in Mali. Providing children with access to primary healthcare is an important development goal. Yet the subsidies needed to achieve this may lead to inefficient overuse, particularly if parents have difficulty assessing their child's need for care. For the treatment of acute illness, price elasticities cannot be used to determine welfare effects, because they do not provide information on whether care is used effectively, which in turn depends on when it is sought. We propose a dynamic model of healthcare timing and define over- and underuse as seeking care too early or too late during an illness spell. We then use nine weeks of daily health records to identify misuse in our sample relative to WHO standards of care. Hazard estimates of care seeking show substantial underuse, but almost no overuse in our population. The primary barrier to the optimal timing of care seeking is cost, not information: subsidies increase care seeking by about 250%, and only 18% of this increase constitutes overuse. In contrast, healthworkers do little to reduce (already minimal) overuse, and may increase underuse when not paired with free care, as we predict in our dynamic model. Free care increases the value of care consumed without crowding out private spending, and it reduces mothers' concern and average illness duration.

*Brown University, corresponding author, email: anja_sautmann@brown.edu.

†Brown University, email: samuel_brown@brown.edu.

‡Columbia University, email: mark.dean@columbia.edu.

We have received invaluable advice and support from Andrew Foster, Duncan Thomas, and Ira Wilson. We would also like to thank Anna Aizer, Dan Björkegren, Caitlin Cohen, Seydou Doumbia, Chris Gill, Jessica Goldberg, Bentley McLeod, Emily Oster, Simone Schaner, Dan Silverman, and Dean Young, and participants of the "Natural Experiments and Controlled Field Studies" workshop 2015. Special thanks go to Hamadoun Bocoum, Zorina Curry, Judith Kom, and Fred Wensing, Mali Health and its staff and Board, and Innovations for Poverty Action. We are grateful for funding from Brown University through the Brown Seed Fund Award, the Population Studies and Training Center, and the Rhodes Center, from the Aga Khan Foundation, and from the ESRC/DFID Development Frontiers Award (ES/K01207X/1). All errors are ours.

1 Introduction

Global health outcomes have improved greatly between 1990 and 2015. But even though mortality of children under five more than halved, six million children died in 2015 – around 16,000 per day. An important contributory factor to child mortality is lack of adequate primary care. For example, in 2013, there were 411,000 malaria deaths amongst children under five in Africa alone, yet less than 26% of children with malaria are estimated to have received adequate treatment. Prior research suggests that improved primary health coverage could prevent 29%–40% of post-neonatal deaths in developing country contexts.[1]

How best to provide primary healthcare for children in developing countries is subject to a longstanding debate, due to the unresolved tension between the goals of providing universal access to care and preventing unnecessary spending that does not actually improve health outcomes. The policy controversy has centered on whether healthcare should be funded via fees paid by patients at the point of use, at the risk of restricting its availability to the poor, or via outside subsidies, which may lead patients to overuse free services and waste precious resources that are needed elsewhere.[2]

Much of the existing literature on healthcare subsidies has focused on calculating the price elasticity of demand for health services, motivated by a static model of healthcare demand. Yet in the context of acute care, an identical increase in utilization could mean that sick children are getting timely access to desperately needed care, or that healthy children are being dangerously overmedicated. In order to judge welfare effects, we need to be able to discern a desirable decrease in underuse from an undesirable increase in overuse, and to do so, information on the timing of care is crucial. When a child gets sick, the longer the symptoms persist the more likely it becomes that they are caused by a serious illness that requires care. Thus, the healthcare decision involves a trade-off between immediate, but possibly unnecessary care, and waiting to see if care is really needed. The choice is thus not just if, but *when* to see a doctor during an illness.

In this paper we assess the impact of healthcare subsidies, and the complementary policy of improved parental information, on the timing of children's care conditional

[1] United Nations Inter-Agency and Expert Group on MDG Indicators [2015], WHO Global Malaria Programme [2015], Bhutta et al. [2008].

[2] See Akin et al. [1987], Litvack and Bodart [1993], Jimenez [1995], UK Secretary of State for International Development [2009], Ridde and Morestin [2012]. A recent literature has addressed the role of prices in determining allocative efficiency for health resources more broadly - for example Dupas and Cohen [2010], Ashraf et al. [2010].

on health status. We model the household's dynamic decision *when* to seek health care during an illness spell in order to define over- and underuse of acute care and to understand the roles of information and subsidies for healthcare decisions. We then estimate the impact of subsidies (which remove cost barriers) and healthworker visits (which remove informational barriers) on healthcare usage with data from a randomized control trial (RCT). Nine weeks of uniquely detailed, diary-style health data on symptoms and healthcare visits from over 1500 children in Bamako, Mali, allow us to observe parents' choice of *when* to seek care in an illness spell and to classify illness days by medically defined need for care following standards of the World Health Organization (WHO). By comparing parents' decisions to these standards we can identify the impact of subsidies and information on over- and underuse.

We model the choice to seek care for a child as the result of a dynamic decision in response to a negative health shock. On every day of an illness spell, the parent has to choose between visiting a doctor and receiving treatment, or delaying the decision for another day. Some delay may be optimal if there is uncertainty about the illness the child has: initially, symptoms may be caused by an illness that will quickly pass on its own, which would render a costly doctor visit unnecessary, but as symptoms wear on, this is less and less likely. The optimal time to seek care then depends on the parents' learning process about the child's recovery probability, as well as the costs and benefits of care.

From the policymaker's point of view (e.g. the WHO), parents may overuse care by seeing a doctor *too early*, on a day when the chance of spontaneous recovery is still high enough to warrant the disutility of waiting an additional day. Underuse occurs if parents seek care *too late* relative to the policymaker's preference. This disagreement may arise for two key reasons. First, parents may have a different assessment of the cost and benefits of care, for example if they are credit-constrained, fail to account for externalities from preventing the spread of infectious disease, or do not face the full social cost of their treatment. Second, unlike the policymaker, parents may be unable to discern illnesses with different chances of unaided recovery, so that they seek care too late when the true probability of recovery is low, and too early when it is high.[3]

In this set-up, we show that subsidy and information policies complement each other in reducing over- and underuse. Subsidies alone can reduce underuse, but may

[3]We discuss the issue of potentially low quality of care and the relationship with underuse in section 3.1.

at the same time induce significant overuse if parents cannot distinguish harmless from serious illness. Better information can help parents differentiate between high and low-severity illnesses, which has the potential to reduce misuse. However, if parents disagree with the policymaker on the net benefit of care, better information may enable them to rule out the most serious cases of illness and therefore paradoxically act to *increase* underuse. When combined, healthworker visits ensure that parents base their decision on good information, while subsidies align their decision rule (conditional on information) with the policymaker's.

In order to study these effects, we collaborated with the NGO Mali Health to carry out an RCT of their Action for Health program, which removes (certain) user fees (the "free care" treatment) and provides free community healthworker visits (the "healthworker" treatment) for children under five years of age. We randomly assigned our sample population to one of three treatment groups – healthworker visits, free care, or both combined – and a control group. The free care arm was provided through the two main local public clinics that collaborate with Mali Health. The healthworkers improve information by checking the child's health and encouraging the use of formal healthcare according to WHO recommendations. This includes accompanying an ill child to the doctor, but also recommending home remedies for ailments that do not (yet) require medical attention.

Six months after program roll-out, we carried out a nine-week panel survey of the 1532 study children, collected weekly in order to overcome recall bias (Das et al. [2011]). The data contains information on 14 symptoms, which enables us to group illness days into spells and classify each spell day as 'care required' or 'early' for care according to WHO care seeking guidelines, yielding information on 3160 illness spells.[4] Our detailed health consultation and treatment data tell us when during a spell a doctor visit occurs and what the costs of the visit are. We use this data to estimate the daily hazard of seeking care conditional on whether care is required by WHO standards, and how this hazard is affected by the policy interventions. In addition, we study effects on the cost of care and on some health outcomes.

Our hazard analysis yields three central results. First, without subsidies, there is extensive underuse and almost no overuse in our study population. 10% of spells that enter 'care required' status actually receive care, compared to only 2.1% of spells which end on an 'early' day. The reason for the difference is that families can discern when

[4]These same guidelines are also used in Mali Health's healthworker training, see below.

care is not yet needed: the probability of seeking care on an 'early' day is only 15% of that on a 'care required' day, and very close to zero.

The second result is that free care dramatically decreases underuse, with almost no increase in overuse. In the free care treatment, the probability that care will be sought on a given illness day increases by 250%. The additional visits are overwhelmingly concentrated on 'care required' days, so that the proportion of care-required spells that receive care rises to nearly 30%, but that of early spells to only 7%. Approximately 82% of additional doctor visits due to free care occur on care-required days. Despite the positive effect of the subsidy, there is still substantial underuse, likely due to the remaining financial costs to care (discussed below) as well as the non-pecuniary costs of a healthcare visit (for example the time required to take the child).

Our third finding is that, contrary to the intentions of the policymaker, healthworker visits on their own do not increase and in some specifications significantly decrease healthcare utilization, and they do so *more* on 'care required' days. In combination with free care, healthworkers have little overall effect on care seeking.

Together, these findings suggest that the primary barrier to timely use of care in this population is cost, not information. There is significant underuse even with subsidies. Parents are able to identify serious illness in their children, as indicated by the differential treatment rates on 'early' and 'care required' days. When combined with free care, improved information has little impact on over- and underuse, partly because there is not much overuse to curb. However, when care is not free, parents appear to make use of the new information in ways that lead them to seek care *less*. One reason may be that families learn from the healthworkers when the child is not in immediate danger. These results are in line with our theoretical predictions for the case where parents and policymaker disagree substantially on the cost threshold for seeking care. T his disagreement is only partially alleviated by subsidies.

Our data and approach highlight the importance of appropriately accounting for censoring due to spontaneous recovery. In our sample, more than 85% of our illness spells end without the use of formal care. This is a form of censoring, because the choice of when to seek care is effectively not observed. We show that simple group averages of the number of untreated 'early' and 'care required' days per spell give the misleading impression that the primary effect of free care is to increase overuse, due to the fact that 'care required' days always occur at the end of the illness spell, and so are more likely to be censored. This problem would be compounded if we were

to analyze only uncensored spells, for example from data collected at the healthcare provider. This approach would lead us to conclude that subsidies *delay* care, due to the selection of families into care who tend to wait longer. Lastly, while our hazard analysis shows that the probability of visiting a doctor on a given 'early' day is nearly zero, the high incidence of *underuse* means that overuse measured at the provider appears more serious, because 17% of consultations occur on early days. Thus, these alternative approaches would lead us to conclude that free care leads to greater inefficiency in care-seeking, in the form of frivolous overuse and potentially detrimental delays until care is sought, than is really the case.

To complement our utilization results, we also study the effects of our policies on private and social healthcare costs, and on health outcomes. We find that free care roughly doubles the value of care consumed, with remarkably little crowding out of private spending on healthcare. This means that families face significant financial costs when they visit the clinic even in the free care group. There is no evidence that clinics raise treatment costs in response to the subsidy. Our study is not powered to detect long term health improvements, and indeed we do not find a measurable effect of the program on total illness, but we do show that the subsidy program is associated with lower average spell length, and that mothers worry less about their children.

There are rich literatures in both health economics and development economics on the effects of subsidies and information provision on use of care, of which the next section provides a more detailed review. Our study contributes in particular to an emerging literature in development economics on the effects of prices on utilization and on the interaction of information and subsidies (see Kremer and Glennerster [2012]). Especially relevant are Cohen et al. [2015], who study the joint effects of subsidized access to malaria medication and information about health status (in the form of a rapid malaria test) on overmedication, and Ashraf et al. [2013], who have shown that information on a new healthcare product complements subsidies in increasing uptake. The first study is designed for a context in which overuse is the primary concern (purchasing drugs despite a negative test result), and the second focuses on preventive care where overuse is not an issue. Our conceptual framework and household-level data allow us to examine the timing of the first contact between patient and provider, a context where both over- and underuse are potential problems.

The next section reviews the existing literature. Section 3 proposes a dynamic model of healthcare demand, analyzes the causes for over- and underuse of healthcare,

and generates predictions on the effects of the subsidy and information interventions of Action for Health. Section 4 gives an overview over the policy environment and provides a description of the study and survey design. Sections 5 and 6 describe the estimation approach and present the main estimation results from proportional and non-proportional hazard models of healthcare demand. Finally, section 7 discusses private and social cost changes as well as health outcome changes as a result of the demand change induced by the subsidy arm of the Action for Health program. Section 8 concludes.

2 Related Literature

Our work contributes to a range of literatures in health and development economics.

The workhorse model of healthcare demand in development economics assumes a health production function that translates health expenses into a health capital stock (see Akin et al. [1986], Foster [1995], Gertler et al. [1987], Gertler and Van der Gaag [1990], Sahn et al. [2003] for examples). This approach is well-suited to modeling the acquisition of human capital, but less so to studying the demand for (curative) healthcare in detail. We propose a new, dynamic model of the choice when to seek care as a function of information and prices. With the exception of Gilleskie [1998], few other studies have explicitly modeled this decision process.

An important reason for this is that most surveys lack the data that would be necessary to successfully estimate such a model. Many seminal contributions in the health economics literature on the effect of pricing on healthcare usage rely on administrative data, like vital registration records or insurance claims (see e.g. the RAND insurance experiment, Manning et al. [1987], and the Oregon experiment, Finkelstein et al. [2012]).[5] Illness spells in this context can only be studied indirectly by grouping insurance claims data into treatment episodes.[6] Other approaches attempt to deal with unobserved health shocks and the resulting censoring of health expenditure at zero using so-called two-part models or selection models, which account separately for the

[5]Recent work in this context focuses on specific data challenges like censoring and endogeneity, e.g. Kowalski [forthcoming]. Cabral [2012] and Aron-Dine et al. [2015] study a specific aspect of demand timing, namely the strategic response to nonlinear insurance plan pricing in the US.

[6]See Stoddart and Barer [1981], Hornbeck et al. [1985], Keeler et al. [1988], Santos Silva and Windmeijer [2001]. Note, however, that there is a public health literature which studies the covariates of the delay in seeking care for specific diseases (see Storla et al. [2008] and Nguyen et al. [2010] for reviews).

decision to use care (based on unobserved health status) and the amount spent conditional on seeking care (see e.g. Jones [2000]). Early research in development economics, instead of estimating demand directly, has focused on the effects of user fee changes on health, labor market participation, and income (Gertler et al. [1987], Gertler and Van der Gaag [1990], Dow et al. [2001], Nabyonga et al. [2005], McIntyre et al. [2006]).

By contrast, our data allows us to to model and estimate the hazard of primary (curative) care conditional on health status and illness duration, and to assess over- and underuse of care relative to the benchmark of WHO care recommendations. We observe illness spells at the household, even if they do not lead to healthcare use, and we show that the selection problems that would occur otherwise are considerable. The one other day-to-day spell data set we are aware of, the 1987 National Medical Expenditure Survey, collects year-long health logs, but it only includes spells that lead to either medical care use or work absenteeism. Gilleskie [1998] estimates a structural dynamic choice model that attempts to address the selection problem by estimating the probabilities of unobserved health events parametrically.

Our study also adds to a recent strand of literature in development that uses field experiments to study the uptake and utilization of healthcare.[7] Several studies address utilization of specific products in response to price changes and information interventions. Ashraf et al. [2013] show that better information about the properties of a preventive water disinfection product can increase the purchase response to a subsidy (the study does not directly address over- or underuse). In related work, Ashraf et al. [2010], Dupas and Cohen [2010], Dupas [2014] and Fischer et al. [2014] examine the utilization and long-run uptake of healthcare products in response to subsidies and test for behavioral "sunk cost" or reference dependence effects.

Most relevant to the present study is Cohen et al. [2015] (henceforth CDS), who examine the effect of subsidies for malaria combination therapy (ACTs) and information about health status in the form of subsidized malaria tests. CDS are most interested in the risk of overtreatment with low-cost over-the-counter medications. They therefore model the (static) decision whether to buy a particular medication, and only carry out malaria tests in the subsidy treatment groups at the point of purchase. In contrast with our results, they find significant rates of overtreatment, which increase with the

[7]More broadly there has been a surge in randomized studies on healthcare use and quality and the resulting health effects. Examples are Miguel and Kremer [2004], Das et al. [2015], Currie et al. [2011], see also Kremer and Glennerster [2012].

rate of subsidization. As in our study, better access to information in the form of subsidized testing does little to change healthcare usage; at least half of the subjects who test negatively for malaria still purchase the ACT. Because of their focus, CDS do not study the universe of illness episodes and treatments, but focus on malaria and ACTs. They do not model or estimate the timing of healthcare, and their data does not contain household level records that would allow them to identify underuse directly. Instead this has to be inferred from auxiliary data.

Our work is also closely related to a set of papers which report the results of randomized trials on health insurance subsidies. We are aware of three interventions, in Nicaragua, Mexico, and Ghana (Thornton et al. [2010], King et al. [2009], Powell-Jackson et al. [2014]). The programs in Latin America had few measurable health or utilization effects, and uptake and retention were low. The trial in Ghana shows some increase in average utilization, but no overall effects on a number of health indicators, and the price elasticity of healthcare demand appears to be small; however, there are significant positive health outcome effects in the subpopulation of children who were anemic at baseline. None of these studies examine the impact of subsidies on over- and underutilization, but instead, they focus on objective health measures which may indicate improved healthcare allocation indirectly. A possible explanation for the low demand response observed in these RCTs, given that the need for acute care is conditional on negative health shocks, is simply that the populations studied are relatively healthy. Another explanation is that they do not use healthcare even when sick, which would mean that there is underuse that the interventions do not remedy. Our study highlights that we cannot know which is the case without measuring healthcare demand conditional on health status.

3 A Dynamic Model of Demand for Healthcare

In this section we develop a model of the demand for acute healthcare to analyze the impact of subsidies and information policies, and to define what constitutes "overuse" or "underuse" from the perspective of a policymaker.

Our model is based on two core ideas. First, demand for acute care is latent, because parents will not seek care for a healthy child.[8] Care is thus only observed in

[8]The implicit assumption is that there is some (minimal) effort cost and no benefits to an acute care visit besides relieving illness.

response to a negative health shock that causes discomfort to the child and carries the risk of long-term harm. Second, the decision to get care is a dynamic problem: parents must decide *when* during an illness spell to seek an evaluation by a (formal) healthcare provider. The key tradeoff in this decision is due to the fact that parents learn more about the severity of an illness over time. Initially, unaided recovery is likely, and it is worth tolerating some disutility from sickness while waiting; but after some delay the probability of recovery without treatment decreases, and further suffering can only be avoided by seeking care. The policymaker may find that parents seek care too early or too late relative to her own preferences, either because they have worse information about the recovery probability associated with a given set of symptoms, or because they have different views regarding the costs and benefits of treatment.

3.1 Basic Model

Consider a child in an ongoing illness spell, defined by a set of observed symptoms $\gamma \in \Gamma$, for simplicity assumed unchanged during the spell. Let $t \geq 1$ denote the number of days the child has been sick, including today. Parents experience sickness disutility $-S < 0$ in every period their child is ill. In any period the parent can choose to take their child to the doctor which (again for simplicity) we assume to lead to immediate recovery.[9] The expected cost of a formal consultation in utility terms is given by $C > 0$. These costs comprise expenses associated with the visit, as well as any travel and opportunity costs. A given monetary expense may have a higher utility cost for poor or credit-constrained households.[10]

When the child falls ill, parents may rationally delay seeking medical care if there is learning about the chance of spontaneous recovery. We think of an illness spell with symptoms γ as being caused by one of a set of possible underlying conditions $\{I_1, ...I_N\}$, where each illness I_n differs from the others in its probability of spontaneous recovery $(1 - \pi(I_n))$. Parents' period-t beliefs about which illness their child has are given by

[9]This assumption may be inaccurate if treatment can be unsuccessful, for example due to low quality of care. For the model, this added complexity would change the value functions and introduce the possibility of repeat visits, but we conjecture that the basic insights into parental behavior remain the same. The expected treatment success also matters for the value of a doctor visit and we will discuss this in the comparative statics section 3.3.

[10]Both S and C may be symptom-specific; for ease of notation we suppress the dependence on γ throughout this section. The analysis does not change substantially if illness disutility and doctor costs change during the spell, as long as $\frac{C}{\delta(C+S)}$ is weakly decreasing in t.

$P_t = \{P_t(I_1), ..., P_t(I_N)\}$, and so the expected probability of remaining ill in period t is $\pi(P_t) = \sum_{n=1}^{N} \pi(I_n) P_t(I_n)$.

An immediate consequence of this model is that parents become more concerned about their child's illness the longer it goes on, because the probability of spontaneous recovery decreases over time. This follows from the fact that illnesses with a low recovery rate become relatively more likely as time passes and the child remains sick.

Lemma 1. *The probability of spontaneous recovery declines over time:*

$$1 - \pi(P_t) \geq 1 - \pi(P_{t+1}).$$

It is constant only if $\pi(I_n)$ is equal for all I_n in the support of P_t.

Proofs are relegated to appendix.

With a fixed recovery probability, the choice to seek care is a simple yes/no decision on the first day of illness. But if learning occurs when observing (non-)recovery over time, the choice of visiting a doctor becomes a dynamic problem. Based on their current beliefs, the parent can decide to consult a formal provider today, or delay the decision to the next day. In the case of delay, the child remains ill with probability $\pi(P_t)$, and the parents update their $t+1$ beliefs about the chance of recovery. With probability $1 - \pi(P_t)$ the illness passes and no doctor visit is necessary. The value function of the parent can be written as

$$V(P_t) = \max\{-C,\ \delta\pi(P_t)(-S + V(P_{t+1}))\}$$

where future costs are discounted by δ.

A solution to this decision problem is based on a cutoff strategy in beliefs.

Proposition 1. *An optimal strategy for the parent is to seek formal healthcare if and only if*

$$\pi(P_t) \geq \frac{C}{\delta(S + C)} = K.$$

In other words, parents wait until the probability of remaining ill rises above some threshold K before seeking care. This threshold is increasing in the utility cost of treatment, and decreasing in the benefit of treatment and the discount factor. The simple form of this solution follows from the fact that parents' beliefs about the 'severity' of the illness monotonically increase over time.

3.2 Comparison with the Social Planner

We can now compare the optimal decision of a policymaker to that of the parents. In our setting, the policymaker is the NGO, who follows WHO and UNICEF guidelines for care seeking (see section 5 below).

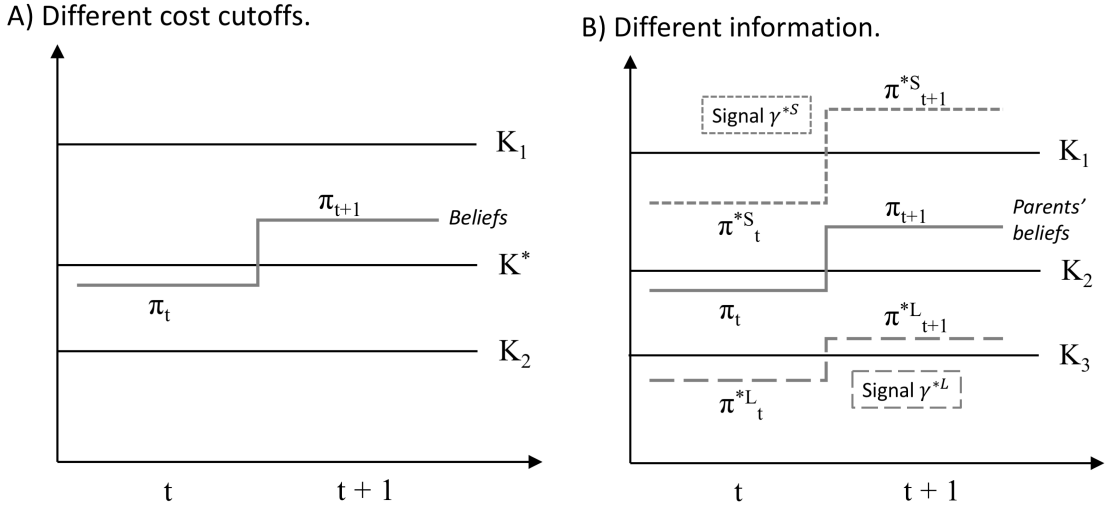
We assume that the policymaker solves a problem similar in structure to that of the parent, but with different parameters. First, the policymaker may be better able to interpret a particular set of symptoms. In order to capture this, we assume that, while the parent observes symptoms γ, the policymaker can differentiate between a number of sub-cases $\{\gamma^{*i}\}_{i=1}^{M}$ with distinct distributions over possible illnesses (though with the beliefs of the policymaker and parent connected via Bayes' rule). We denote as π_t^{*i} the beliefs of the policymaker at time t after having observed symptoms γ^{*i}.

Second, the policymaker uses discount factor δ^*, social costs of care C^*, and disutility of illness S^* to establish their cost threshold K^*. There are many reasons why policymaker and parent may disagree on these parameters. First, the policymaker may apply different costs to care. For example, if a parent is extremely poor and credit-constrained, the utility cost of taking the child to the doctor may be very high (perhaps because it means the family cannot eat that day), while the policymaker might price the visit at a lower rate (perhaps based on the opportunity cost of treating other children). Second, parents may evaluate the benefits differently; for example, they may not account for positive externalities to treatment due to the risk of infection from a sick child. Parents may also disagree with the policymaker on the value the child's (next-period) discomfort through S or δ. Lastly, since healthcare is an experience good, parents may *ex ante* misjudge the value of S, for example because they believe the quality of care to be low. Note that a probability of unsuccessful treatment can be modeled by replacing S with the increase in the (expected) flow utility after treatment;[11] this does not affect the comparative statics unless policymaker and parents *disagree* on this probability.

For the remainder of the paper, we define over- and underuse in relation to the policymaker's preferences. "Overuse" is equivalent to the parents seeking care too early: the policymaker considers the chance of spontaneous recovery high enough to warrant the disutility of waiting when the parent does not. "Underuse" occurs when parents seek care too late: to the policymaker, a failure to seek care only delays a doctor visit

[11] But note that our model does not allow for repeat consultations after unsuccessful treatment, see earlier remark.

Figure 3.1: Policymaker vs. parental optimum.



that will eventually occur with high probability anyway. If the disagreement between the parents and the policymaker is due to externalities or incorrect beliefs, e.g. about recovery probabilities or the benefits of formal care, adherence to the policymaker's preferences will unambiguously improve social welfare. However, at prices that reflect marginal cost and with full information, parents' choices may be efficient even if they constitute over- or underuse according to the policymaker.

We demonstrate the causes of over- and underuse with two examples. The first shows that if the policymaker and the parents disagree only about the costs and benefits of care, so that $K \neq K^*$, there will be underuse *or* overuse, but not both. The second shows that overuse and underuse may co-exist if the policymaker is better informed about the underlying illness than the parents are, even if they share the same cost cutoff.

Example 1 (Same Information, Different Cost Cutoff)**.** Assume that symptoms γ can be caused by two equally likely illnesses, one severe (I_S) and one less severe (I_L), with $\pi(I_S) > \pi(I_L)$. Bayesian updating implies that $\pi_t < \pi_{t+1}$ (see appendix A for numerical details). We first consider the case in which the policymaker and the parent have the same information, but different cost cutoffs. Beliefs about recovery are illustrated in Panel A of Figure 3.1 by the solid gray horizontal line. Suppose the policymaker has cost cutoff K^*, and therefore would like the child to go to the doctor if they are still ill in period $t+1$. If the parent has cost cutoff K_1, they will not take the child to the

doctor until after $t+1$ – an incidence of underuse.[12] If instead the parent has cost cutoff K_2, they will take the child in period t – an incidence of overuse. In general, cost differences can lead to underuse (if $K^* < K$) or overuse (if $K^* > K$), but not both.

Example 2 (Same Cost Cutoff, Different Information)**.** Now suppose that the policymaker can interpret the symptoms better than the parent. Specifically, she can distinguish two equally likely sets of symptoms γ^{*L} and γ^{*S}. Symptoms γ^{*L} indicate lower severity, while with γ^{*S} the child is more likely to be seriously ill. An example is the distinction between different types of diarrhea, where blood in the stool is a sign of dysentery and therefore potentially serious illness.

For concreteness, assume that under γ^{*L} there is an initial probability of illness I_S of 0.25, and a probability of illness I_L of 0.75. It can be shown that $\pi_t^{*L} = \pi(P_t(\gamma^{*L})) < \pi_t$, and $\pi_t^{*L} < \pi_{t+1}^{*L} = \pi(P_{t+1}(\gamma^{*L})) < \pi_t, \pi_{t+1}$. To the policymaker observing the low severity symptom, the probability of not recovering is lower than to the parent in either period. Under γ^{*S}, the probabilities of illnesses I_S and I_L are reversed, so that $\pi_t^{*S} > \pi_t$ and $\pi_{t+1}^{*S} > \pi_{t+1}$. Panel B of Figure 3.1 illustrates the beliefs of the parents who observe γ (solid gray line), and the policymaker who observes symptoms γ^{*L} or γ^{*S} (lower and upper dashed lines, respectively).

Consider first the case in which both the parent and policymaker have cost threshold K_1. The policymaker would like the child to go to the doctor in period $t+1$ if symptoms are γ^{*S}, but not otherwise. However, given the information the parents have, they will not take the child to the doctor in either period, so there is underuse for the more severe set of symptoms. Similarly, if the cost threshold is K_3, the policymaker would like the child to delay the doctor visit until period $t+1$ if the 'true' symptoms are γ^{*L}, yet the parent will always go in period t, so there is overuse. A special case occurs if the cost threshold is K_2. The policymaker prefers a doctor visit in period t if the symptoms are γ^{*S}, but in neither period with symptoms γ^{*L}. The parent, however, will always seek care in period $t+1$: lack of information can *simultaneously* lead to overuse and underuse.

[12] The parent will eventually take their child to the doctor if K_1 is below $\pi(I_L)$, otherwise they will never go.

3.3 Policy Impact

Within our model, we interpret the two interventions of our program as a reduction in the (expected) cost of treatment, and a change in information. The "free care" component of Action for Health reduces the cost of treatment C and therefore the care-seeking threshold to $K' < K$. The healthworkers teach the families which symptoms should trigger a doctor visit according to the policymaker's beliefs (here those of the WHO or Mali Health). Thus, the "healthworker" component aligns the information of the parent with WHO guidelines. In the example above, this would imply that the parents learn to distinguish symptoms γ^{*L} and γ^{*S}. We will assume that, prior to the Action for Health program, parents may value healthcare less than the policymaker (i.e. use threshold $K \geq K^*$) and potentially hold uninformed beliefs.

First we consider the impact of healthcare subsidies.

Claim. By lowering the parents' cost threshold to $K' < K$, free care weakly decreases underuse, but may not eliminate it if K' remains above K^*. It may also not eliminate underuse if the parents have worse information that the policymaker.

The first part of the statement is immediate. That underuse may remain with bad information, even if cost thresholds are aligned, follows from Example 2 when the policymaker holds beliefs π^{*S} (Panel B, Figure 3.1).[13] At $K' = K_2 = K^*$, parents with beliefs π seek care in period $t+1$, but the policymaker would seek care even earlier in period t.

Claim. Free care can *increase* overuse if (and only if) it pushes the cost threshold of the parent below that of the policy maker ($K' < K^*$) or if the parents have worse information than the policymaker.

The first case is illustrated by Example 1 when costs are reduced from K_1 to K_2. The second follows from Example 2 when the policymaker holds belief π^{*L}. At $K' = K_2 = K^*$, parents seek care in period $t+1$, but unlike before, the policymaker would prefer them to wait even longer, so there is overuse.

More generally, example 2 illustrates that under- and overuse cannot be eliminated by changing the cost of care alone if the optimal action of the policymaker differs between π^{*S} and π^{*L}. This is because uninformed parents cannot discern these two situations and make the same decisions for γ^{*L} and γ^{*S} .

Next we consider the effect of healthcare visits.

[13] We use π to denote the sequence of beliefs $\{\pi_t, \pi_{t+1}, ..\}$ and similarly π^{*S} and π^{*L}.

Claim. Information that aligns the beliefs of the parent and the policymaker may reduce both overuse and underuse. If the cost threshold of the parent is weakly higher than that of the policymaker ($K \geq K^*$), it *must* weakly reduce overuse; if the cost threshold of the parent is weakly lower than that of the policymaker ($K \leq K^*$) it *must* weakly reduce underuse.

This follows because the parent will seek care weakly less often than the policymaker if $K \geq K^*$ and more often if $K \leq K^*$, provided they have the same beliefs.

Claim. Information may *increase* underuse if $K > K^*$.

This can be seen in Panel B of Figure 3.1. Assume that the policymaker has cutoff K_3 while the parent has cutoff K_2. If the parent cannot distinguish between γ^{*S} and γ^{*L}, they always seek care in period $t+1$ (but not in t). This is aligned with the preferences of the policymaker in period $t+1$, although there is underuse in period t if the symptom is γ^{*S}. However, an informed parent will only seek care in period $t+1$ and if they observe symptom γ^{*S}, leading to greater underuse, here in the case of symptom γ^{*L}. Teaching parents to tell apart severe cases of illness, e.g. by looking for blood in the stool, means that they also learn to distinguish less severe cases. This can be detrimental if it leads to underuse in those cases.

3.4 Interpreting the Results of the RCT

We can now make predictions about the impact of the policies of the randomized controlled trial when there is misalignment between the parents' and policymaker's cost thresholds, their beliefs, or both.

Cost differences only: We expect underuse, but no overuse in the control (assuming $K \geq K^*$). Free care will reduce underuse, and if it pushes the cost threshold of the parent below that of the policymaker, it will eradicate underuse but cause overuse. Healthworkers have no effect.

Information differences only: If cost thresholds are aligned, but parents have limited information, both underuse and overuse occur in the control. Parents exhibit similar care seeking behavior on early and care-required days. Healthworker visits will reduce both overuse and underuse, while free care will reduce underuse but also (substantially) increase overuse.

Information and cost differences: Underuse is prevalent in the control, but there may be some overuse. Healthworker visits alone may decrease overuse but also *in-*

crease underuse. Free care alone will reduce underuse but increase overuse at the same time. Free care and health worker visits combined can simultaneously reduce over- and underuse. Thus, information and subsidies are complements that can simultaneously address both types of misallocation.

Two sources of heterogeneity may blur these predictions. First, there may be heterogeneity in cost cutoffs and in the level of information that parents have initially. Second, the policies may have different effects for illnesses with different associated beliefs and costs of care. Heterogeneity in cost cutoffs can lead to both over- and underuse in the population even when parents hold correct beliefs. Different sets of symptoms γ may lead to each of the situations described in example 2; lack of information may cause underuse, overuse, or both for different illnesses. The comparative statics continue to hold, however: if information is the important constraint, then healthworkers should have an impact on both over- and underuse, while free care should reduce underuse, possibly at the cost of higher overuse. If cost is the constraint, healthworkers should have little effect while free care reduces underuse only, unless the subsidy lowers parents' K below K^*.

4 Study Background and Data Collection

The Malian healthcare system builds on a network of community health clinics or *centres de santé communautaires* (CSCOMs). A CSCOM has typically one primary care provider on duty, along with a handful of other staff, and sells prescribed medications through an attached pharmacy that stocks basic supplies and treatment for common illnesses. Public care is partly subsidized by the government and by private NGOs, but primarily funded by user fees, in accord with the Bamako Consensus from 1981 which advocates self-sustaining, decentralized primary healthcare in West Africa. The public healthcare system is flanked by a private formal sector and a large informal sector, consisting of itinerant peddlers and market stalls that sell "Western" drugs, traditional healers, and marabouts or Islamic healers.

This study was conducted in Sikoro, a peri-urban region in a hilly area on the outskirts of the capital, Bamako. The majority of its roads are unpaved, and most dwellings are not connected to the water supply or the sewage system. During the rains from roughly July to October, the incidence of infectious disease is highest, due to diarrhea, malaria, and respiratory infection. Mali is one of the poorest countries in the

world, and especially rural areas have very high rates of maternal and child mortality, but poor urban areas also often lack basic health services, despite better accessibility. Peri-urban populations as in Sikoro are one of the fastest-growing demographics in West Africa.

4.1 Study Design

Mali Health started their Action for Health (AfH) program in 2010 in collaboration with the two local CSCOMs in Sikoro. The CSCOMs in this study were subject to basic quality control by Mali Health.[14] The AfH program combines healthworker visits and subsidized care. The healthworker component provides biweekly visits from community healthworkers (CHW). CHWs live in the area and are trained and employed by Mali Health. They monitor children's health, advise families on acute care, and most importantly encourage them to see a formal doctor in the event of illness. They are trained following the C-IMCI (Rosales and Weinhauer [2003]), a set of guidelines for community healthworkers that incorporates the WHO's and UNICEF's "Integrated Management of Childhood Illness" recommendations for when to refer children to formal healthcare (WHO [2014], WHO Department of Child and Adolescent Health and Development [2005]; see also below).[15] The free care part of AfH is administered via a personalized card that entitles the child to a free consultation at one of two local CSCOMs and free treatment for any illness due to diarrhea and malnutrition, malaria, vaccine-preventable diseases, and respiratory infection (the five main causes of child mortality). This reduces the average visit cost substantially, but not to zero (see section 7.1). Households are eligible for program enrollment if they pass a basic proxy-means test, designed to select about the poorest third of families in the area.

The research design took advantage of the second planned roll-out wave of AfH in late 2012. Mali Health conducted a census in their new expansion area in summer 2012

[14]Mali Health in particular controlled costs in the free care group (see also section 7). This is important because it keeps the supply side constant across treatment groups. A larger issue is the generally low quality of care in public healthcare systems in developing countries (see e.g. Das [2011], Das and Hammer [2005, 2007, 2014], Das et al. [2015], Das and Sohnesen [2006], Leonard and Masatu [2010], Leonard et al. [2002]). This may affect how desirable the use of the healthcare system is from the policymaker point of view. *Conditional* on the policymaker's preferences we show in the theory section that quality of care affects our model only if parents disagree with the policymaker's quality assessment.

[15]Outside the focus of this paper, CHWs also advise households on preventive care, and dispense water chlorination tablets for households with no access to clean water.

to enumerate all eligible families based on geography, the presence of children under five years of age (or a pregnant mother), and the proxy means test. Data was collected in two survey rounds in 2012 and 2013 in the rainy season. Households identified by the Mali Health census were revisited for the baseline survey in 2012. All households that were found at baseline were included in the random assignment to the different treatment groups.

Data was collected at the level of the household, defined as all persons who identify the same individual as their household head.[16] The unit of randomization is the compound. A compound may house more than one family, and typically consists of a few rooms around a common courtyard with shared latrines and other facilities. After stratifying compounds by average household assets, number of eligible children, and location, each was assigned to one of four groups: a full treatment group that received healthworker visits and free care, a free care only group, a healthworker only group, and a control. The healthworkers in the two healthworker treatment groups were trained and managed separately by Mali Health, and no healthworker visited families in both the healthworker-only and the full treatment group, to avoid spillovers on the provider side.[17]

In both survey rounds, mothers with eligible children answered questions on their own children, the most senior mother reported on demographics and household finance, and the household head was interviewed on household assets and income. In this study we use demographic, location, and household asset data from the baseline survey, and

[16]The surveyors were instructed to first approach the most senior mother in the compound enumerated by Mali Health. As part of the roster, she is asked whom she considers the head of her family. All in the compound under the same household head are part of the same household. If the household head she names is absent, the mother, her children, and all who consider *her* the de-facto decision maker in family matters constitute a separate household. If there are other eligible women and children in another household in the compound, the surveyor proceeds to the next most senior mother who was not yet included in the survey. This definition aims to identify the respondent who makes decisions over small children's health. The split of a household with an absent head into sub-households accounts for polygamous households where each mother and her children form a "decision unit". This occurs only very infrequently in our data.

[17]Healthworkers could not be assigned across the entire intervention area in order to keep their travel distances manageable. Instead, Mali Health paired healthworkers of similar experience and quality, and one of each pair was assigned to the healthworker only or the healthworker and free care group. The pairs were then assigned in an overlapping pattern by stratum. For example, pair 1 and pair 2 were each randomly assigned half of the families in the same stratum, and half of the families in a different stratum, each shared with another pair of healthworkers. In this manner, each stratum was assigned four healthworkers (two in each treatment group), the quality of the healthworkers in the treatment groups was matched, and no two strata had identical sets of healthworkers.

Table 1: Children Per Group and Attrition.

	Control	Healthworker	Free care	HW & FC	All
Original Sample	465	438	453	418	1774
Not Found at Baseline	30	24	19	13	86
Moved Post-Baseline	34	23	35	23	115
Died Post-Baseline	0	1	4	1	6
Refused Post-Baseline	0	0	0	1	1
Unexplained Absence	6	10	11	7	34
2013 Sample	395	380	384	373	1532
Total Attrition	15.1%	13.2%	15.2%	10.8%	13.6%
Attrition Post Baseline	9.2%	8.2%	11.5%	7.9%	9.2%

daily health diary data on children from the follow-up round.

4.2 The Sample Population

Table 1 shows the size of the sample over time. Mali Health originally identified 1774 eligible children in the census. The baseline survey took place around three months after the initial census. At that point, a total of 86 children were not found by our surveyors. The roll-out of the Action for Health program started another three months after the first survey in early 2013. By the second survey round, an additional 156 children were not found. Of those, we know that six had passed away, and 115 had moved out of the survey area. One family refused participation. In total, the sample in 2013 consists of 1532 children present in both survey rounds, from a total of 957 households and 592 compounds. This corresponds to an attrition rate of 9.2% between baseline survey (and assignment to treatment group) and follow-up. There is no obvious differential attrition by treatment.[18]

Table 2 provides an overview of the demographics and finances of the households in our sample. There are on average more than six people in a household, due in part to polygamous marriages and multi-generational households. Most households are headed by a man. About half of the household heads are literate and two thirds speak the lingua franca of the region, Bambara. Average household assets are about CFA 3 million or USD 6000, with an owner occupancy rate of 43%. Average weekly income is about $61, though only 12% of household heads earn a regular salary (this is in part due to the selection criteria applied by Mali Health).

[18]In addition, we have 2013 data on children who were not present in the 2012 baseline, in particular infants born after the 2012 census. This data is not used in our analysis.

Table 2: Sample and Balance.

	Mean at baseline			Means of baseline variables for households included in 2013.			
				Control	Healthworker	Free care	HW & FC
	Mean	SD	N	Mean	Mean	Mean	Mean
Household Variables							
Household Size (# Members)	6.26	3.14	955	6.15	6.13	6.32	6.44
Assets (USD)	6162	12897	935	5233	6299	7512	5722
Weekly Income (USD)	63.35	76.68	895	58.40	62.67	66.19	66.62
House Size (# of Rooms)	2.06	1.56	931	1.98	2.11	2.14	2.04
Owner Occupier	0.45	0.50	935	0.42	0.47	0.48	0.44
Household Head Variables							
Male	0.89	0.32	962	0.90	0.88	0.88	0.89
Literate	0.48	0.50	963	0.46	0.44	0.47	0.55
Bambara 1st language	0.68	0.47	963	0.73	0.69	0.68	0.64
Salaried	0.12	0.33	963	0.12	0.13	0.13	0.10
Child Level Variables							
Age (Months)	40.21	0.47	1532	40.61	40.47	39.51	40.26
Male	0.51	0.01	1530	0.50	0.48	0.52	0.52
Adopted	0.05	0.01	1468	0.06	0.06	0.06	0.03
Weight for Height Z-Score	-0.61	0.05	1293	-0.65	-0.62	-0.55	-0.63

No significant differences in group means using ANOVA.

The average age of the children in our sample was about 3.5 years at the time of the 2013 survey, reflecting the fact that we did not add children born after the 2012 baseline. Our study children are more than half a standard deviation below their age-typical weight for height relative to the WHO reference distribution. An ANOVA analysis shows that there are no significant differences between the treatment groups for any of the 13 variables listed in table 2.

4.3 Health Calendar

The core element of the data collection is a detailed health calendar for each child, containing information on any symptoms the child exhibited, all consultations that occurred during the week with respect to the child's health, and all medications taken. Health calendar data was collected on a weekly basis over the course of nine weeks from the child's primary caretaker (usually the mother). To aid mothers' memory, they were given pictorial diaries with images representing the different symptoms (Appendix B, figure B.1). They were asked to mark off any symptoms and health-related events on the day they occurred, and the surveyors reconstructed the child's health history together with the mother during the visit following each week. This method provides us with a uniquely detailed record of all the health events involving the study children, without the problems associated with long recall periods (Das et al. [2011]).

4.3.1 Symptom Calendar

The list of symptoms was designed in collaboration with Mali Health staff and based on the C-IMCI. The C-IMCI is designed for use by community healthworkers (CHWs) who have no prior medical experience, and is used in Mali Health's own healthworker training. It consists of simple rule charts involving easy-to-spot symptoms to broadly classify a child's illness and establish need for care. These properties make it well-suited for use in this study: first, they allow the use of symptom reports from mothers and surveyors who are not medically trained, and second, the treatment rules allow us to map observed symptoms into when a child should seek care according to the policymaker.

The symptom survey explicitly names ten symptoms: convulsions fits, or spasms; lethargy or unconsciousness; inability to drink; vomiting; coughing; difficulty breathing; more than three loose stools; blood in the stool; sunken eyes; and unusually hot (cold) skin. These are, for example, indicators for fever ("skin unusually hot") or diarrhea ("more than three loose stools per day"). Other symptoms are recorded free-form. From those records we constructed four additional symptom groups: skin conditions, cold symptoms, ear pain, and injuries. Appendix B provides a detailed description of how symptoms were collected and gives an overview of the recorded symptom days in table 11.[19] We observe on average 60 days per child, and of those the child exhibited symptoms on 18 days. The most frequent symptoms were cold symptoms, coughing, and unusually hot skin. The section on illness spells below provides more details.

4.3.2 Consultations

At each surveyor visit, mothers are asked if they discussed their child's health with anyone in the previous week, and the surveyors recorded the role or occupation of the person seen, the type of facility, and the cost of the visit. For commonly used sources of health care in the area we have GPS location data, otherwise we collected approximate distances from the respondent's house.[20] For each consultation, we have information

[19]The table shows that the full treatment group experiences or reports on average slightly higher rates of illness compared to the control, and it cannot be said if the increase in symptom reporting is partly an effect of the treatment. However, the difference is driven by injuries and cold symptoms, which are not relevant for the Action for Health program: they do not indicate a need for care according to the IMCI, and Mali Health does not pay for healthcare related to either.

[20]Many consultations in fact occur at home, in particular the Mali Health CHW visits, but also drug sales by peddlers or visits from nurses or healers. Virtually all formal consultations, however, do

on actual treatment received and its price.[21]

We group consultations into formal and informal or other care (see table 3). The nearly 600 formal consultations in our data are provided by public facilities including the CSCOMs, and by private doctors and clinics. For the purposes of our analysis, we focus on formal healthcare as the (only) way to address a medical need for care. Informal sources are peddlers and market sellers, stores, healers, marabouts, or "elders". An informal visit is typically not accompanied by systematic diagnosis and provided by individuals who, to our knowledge, do not have formal medical training and are not bound by the hippocratic oath. Drugs received from informal sellers may be stored inappropriately or expired. While informal care is not *necessarily* inadequate, there is by definition no way to control quality and regulate the market. We therefore assume that universal access to high-quality care can only be guaranteed if demand is first channeled into formal care. We decided to exclude pharmacy visits from the 'formal' category on the basis that pharmacists typically do not diagnose patients. Note that parents in some instances visit several care providers in the same day, for example when they get a prescription from a (formal) doctor and then purchase the medication at a pharmacy or a store. We classify the formality of a consultation by the "most formal" in a set of visit records that are linked in this manner.

Table 3 provides an overview of the number of healthcare consultations by type, and their average cost to the household, including medications purchased in connection with the consultation, in West African CFA. 500 CFA equal approximately one US dollar in market exchange rates. From the table, it can be seen that formal care is substantially more costly than informal care. The private cost of care at the CSCOM is relatively low due in part to the free-care intervention. The cost effects of our policy intervention are discussed further in section 7.

4.4 Illness Spells and Optimal Care

We use the symptom data from the health calendar to construct illness spells. An illness spell is any contiguous sequence of days on which a child exhibits one or more

not occur at home.

[21] In order to facilitate the recording of drugs prescribed, purchased, and taken, mothers were asked to keep the packaging of any drugs their child received. The surveyors could search and fill the drug from a database of brand and generic names of about 300 medications commonly prescribed in Malian pharmacies, clinics, and doctor's offices, or sold by traditional healers, at street stalls, and by itinerant peddlers.

Table 3: Average Cost per Consultation, CFA.

	CSCOM doctor or nurse	Private doctor or nurse	Other (pharmacy or informal)
Mean	1388	5523	223
SD	2820	5394	555
N	526	72	2342

CSCOM (public) visits include regional public reference hospitals for the CSCOM. Private doctors include all private or non-governmental facilities. Informal visits include sources like traditional healers, midwives, market stalls, itinerant peddlers, neighbors, and imams. Pharmacies are licensed to sell drugs

symptoms, and it ends when the child is symptom-free. The object of interest for much of this paper is the untreated portion of the illness spell, or the "pre-care" spell. A pre-care spell lasts up to the illness spell day on which the first formal health consultation occurs or until recovery, whichever comes first. Our data contains 3160 useable spells (i.e. spells without left-censoring – see section 5) with an average length of 6.5 days. Incidence is remarkably symmetric between the four treatment groups; almost exactly a quarter of all spells occurs in each of the treatment groups (between 781 and 798 spells, see table 8 below). This is consistent with the fact that the main sources of illness – malaria, respiratory, and gastrointestinal – are infectious and that access to primary care cannot reduce infection rates.[22]

In order to identify over- and underuse on the part of the parent, we need to determine when care should be sought from the point of view of the policymaker. We do this using the C-IMCI guidelines, which describe recommended care, tailored to developing-country settings.[23] We assume that the IMCI guidelines represent the policy aims of the Action for Health program, as well as those of the WHO. Based on the symptoms reported by the mother, these can be used to classify, within each pre-care spell, when a child should receive formal medical care. For example, a child with simple diarrhea is advised to seek care only if the symptoms last for five days. However, if the diarrhea is accompanied by blood in the stool, care should be sought immediately. Every day in the pre-care spell can thus be classified as either an 'early' day (i.e. occurring before care should be sought according to the guidelines) or a 'care required' day (on or after the day at which care should be sought). Appendix B

[22]One might wonder why the healthworkers had no effects on prevention. There are two likely reasons, first, prevention for malaria is already high – a majority of households own and use malaria bed nets. Second, preventive measures for diarrhea (in particular water disinfection) were not successful: in other work we document that families in the healthworker treatments use chlorine tablets significantly more, but that measured chlorine levels in their drinking water are virtually unaffected.

[23]Note that this tailoring should account for both tighter budget constraints as well as lower quality of care than in richer countries.

Table 4: Pre-care days by need for care and symptoms per child.

	Total untreated days per child		Care not required (early)		Care required	
	Mean	SD	Mean	SD	Mean	SD
Number of days:	15.93	(14.52)	7.55	(8.27)	8.38	(11.54)
Percentage where each symptom is present:						
Convulsions, fits, or spasms	0.36%		0%		0.53%	
Lethargic or unconscious	4.06%		0%		8.02%	
Unable to drink or breastfeed	1.08%		0%		2.06%	
Vomiting everything	5.08%		0%		10.93%	
Coughing	32.47%		28.68%		36.48%	
Difficulty breathing	4.54%		3.93%		5.34%	
> 3 loose stools	7.47%		6.66%		10.17%	
Blood in the stool	0.60%		0.05%		1.10%	
Sunken eyes	2.05%		0.16%		3.60%	
Unusually hot skin	31.62%		0.15%		62.49%	
Other: rash, spots, or itch	2.97%		3.49%		1.67%	
Other: cold symptoms	51.11%		59.24%		42.71%	
Other: ear ache	1.00%		0.85%		1.57%	
Other: wound, injury, or burn	4.08%		5.98%		1.74%	
Other symptoms	5.56%		7.22%		2.87%	

Untreated days include left-censored spells.

describes the classification in detail.

The top of table 4 shows the untreated days per child, both in total and split into early and care-required days. Below is the percentage of those days on which each symptom is present. For example, the first four symptoms, convulsions, lethargy, inability to drink, and vomiting, are acute danger signs and should trigger immediate care. On average, nearly 16 (27%) of 60 observed days per child were untreated (pre-care) illness days, and more than half of those indicated an (unfilled) need for care. The biggest contributors to need for care are gastrointestinal symptoms (diarrhea and vomiting) and fever (unusually hot skin). While cold symptoms and coughs do not in themselves indicate a need for care, they co-occur on many "care required" days.

Note that the symptoms we selected for collection do not cover all possible illness, but rather focus on conditions that both mothers and surveyors could easily recognize and report. As an example, one symptom mentioned in the C-IMCI as an indicator of respiratory illness, the number of breaths per minute, was judged to require too much experience and training for the surveyors to be collected reliably. Note also that wounds and injuries are not mentioned in the C-IMCI and were therefore not used to determine need for care. This means that our classification as "care required" represents a lower bound for the true need for care. On some days that we designate as "early" the child may actually require care according to the (full) IMCI guidelines. Furthermore, parents may have private information that indicates a need for care

according to policymaker preferences, even if this is not specified in the IMCI (e.g. if the child has blood poisoning).

5 Estimating Exit into Care

Our aim is to estimate the degree of over and underuse of care by our target population, and the impact of free care and healthworker visits on misuse. In other words, we wish to understand the probability that a parent seeks care conditional on need for care and on treatment assignment. The stylized model of section 3 illustrates two potential challenges for this estimation. First, over- and underuse cannot be measured on an aggregate level, but need to be evaluated day by day during an ongoing illness spell. Second, time-to-care observations are generally censored. Spontaneous recovery can occur, and in those cases we do not observe what parents would have decided if the child had *not* recovered. This type of censoring is important when a large number of spells do not end in formal care. Indeed, only 494 of the 3564 spells in our data set are uncensored.

Censoring due to spontaneous recovery leads to bias if not corrected in the estimation. For instance, the average observed "pre-care" spell length, measured until a formal consultation occurs or until the child recovers, systematically underestimates time-to-care. Moreover, observations with longer time-to-care are more likely to be censored, so the *degree* of censoring is affected by any treatment that changes time to care. For example, suppose the cost threshold K decreases as a consequence of the healthcare subsidy, and time to care falls. There will be spells that are censored at high K – the child recovers before care is sought – but uncensored at low K. The average time to care underestimates spell length at high K by more and will therefore show a smaller reduction in time to care. The same censoring issue would also lead us to misjudge the relative incidence of over- and underuse and their response to the policy intervention, because overuse, in the form of seeking care too early, occurs at the beginning of a spell, while underuse, in the form of not seeking care on a care-required day, can only occur later in the spell. This implies that censoring first affects the care-required portion of the spell, and any policy effects on the number of care-required days will be more strongly attenuated than effects on the number of early days.

These problems are exacerbated with data on uncensored spells only, for example data collected at the healthcare provider rather than the household. In such data,

families with the longest time to care will be least often observed. When treatment costs fall, the families who are newly observed making doctor visits have long time to care compared to those who visit under high cost as well. In the extreme, the number of observations with long delays to care may increase enough to lead us to wrongly conclude that families go to the doctor *later* at lower cost.

We show in Section 6.2 that these concerns are not just theoretical. For example, by comparing simple average spell length and number of days classified as "early" vs. "care required" we would dramatically underestimate the effect of free care on healthcare demand and misjudge its effect on over- and underuse of care.

In order to appropriately deal with the right-censoring issue, we use a hazard model to estimate the probability of receiving formal care on day t of an illness spell, accounting for the fact that exits are not observed if the spell is censored before t.

The hazard of formal care on day t, conditional on the spell continuing, is given by

$$h(t) = \frac{f(t)}{1 - F(t) + f(t)} = \frac{f(t)}{S(t)}$$

where $f(t)$ is the probability of exit at t, and $S(t) = P(T \geq t)$ is the survival function. We estimate the probability of seeking care on a given day in two alternative specifications, a Cox proportional hazard model and a linear non-proportional hazard model.

In a Cox model, the hazard at t conditional on covariates x_{it} is

$$h(t|x_{it}) = \lambda_0(t)e^{x'_{it}\beta}.$$

The Cox model imposes no restrictions on the baseline hazard $\lambda_0(t)$, but assumes that the effect of each covariate shifts the probability of seeking care proportionally on every day t of the illness spell. In a linear, non-proportional model, the constant and the effect of some or all covariates are allowed to vary across t, so that

$$h(t|x_{it}) = x'_{it}\beta_t.$$

Our primary aim is to estimate the degree of over- and underuse of formal care in the study population, and the impact of free care and healthworker visits on misuse. The explanatory variables of interest are therefore (i) indicators for the policymaker's belief about whether care should be sought, (ii) indicators for the free care and healthworker

Table 5: Pre-care spells by type of censoring.

Pre-care spells	Uncensored and right-censored			Left-censored
	Total	Early	Care required	Total
Mean days within spell	6.543	3.216	3.327	10.292
SD	(7.142)	(4.053)	(6.311)	(10.582)
N	3160	3160	3160	404

A pre-care spell ends when a formal consultation occurs. It is censored if no consultation occurs on last day of illness spell or if observation directly before first or after last illness day is missing.

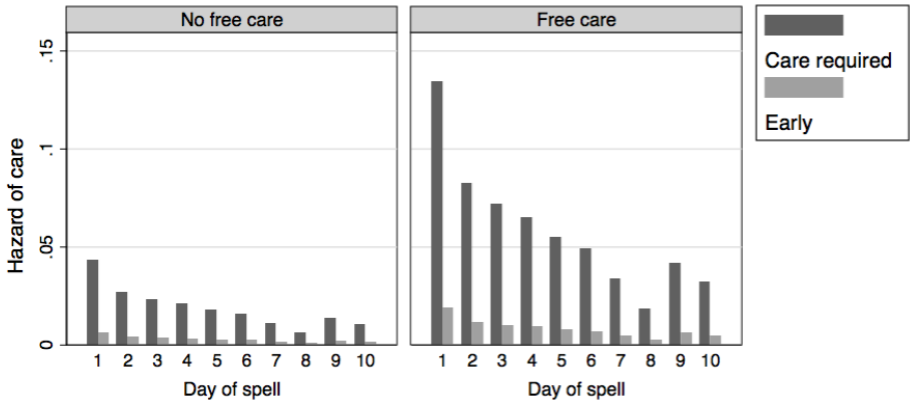
interventions and their interaction (treatment dummies), and (iii) interactions of (i) and (ii). The first tells us to what degree individual behavior agrees with the policymaker's preferences, and if there is underuse of formal care, overuse, or both. (ii) and (iii) tell us how individual behavior responds to the policy interventions. The interaction of the two treatment dummies captures complementarities between information and subsidies. Interacting treatment dummies and policymaker beliefs allows for the demand response to differ depending on the policymaker's preferences. For example, if better information reduces both overuse and underuse, the effect of the healthworker treatment on the hazard of care should be *positive* on days when care is required, but *negative* otherwise.

Table 5 shows some summary statistics on the recorded spells we use for the estimation. The average uncensored or right-censored spell is nearly one week long, with a high standard deviation. Within these spells, there are on average 3.2 "early" days and 3.3 "care-required" days. Due to survey interruptions and spells recorded in the first week of the survey, a number of illness spells are left-censored, that is, they may have started before the first recorded day. Since the day of the spell is unknown for these spells, they cannot be used in the hazard analysis.We check for the robustness of our results by running our hazard analyses treating the left-censored spells as uncensored (see appendix C).

6 Results: The Effects of Subsidies and Information on Demand for Care

Table 6 shows the estimation results from the Cox proportional hazard model. The table reports hazard ratios, that is, the relative increase in the probability of seeking care when the independent variable increases by one unit. The first two specifications in the table show the simple treatment effects of healthworker visits (HW), free care (FC), and the interaction of the two (HWFC, which picks up any additional effect

Figure 6.1: Predicted hazard of formal care from the Cox hazard estimates, by spell day and classification as early vs. care required (for a child with the median sample characteristics and 50% probability of being in the healthworker treatment group, based on specification (3) in table 6). Left without, right with free care.



of combining the two policies). Specifications (3) to (6) add indicators for the IMCI classifications for "early" days where care is not yet required. Models (2), (5), and (6) control for the household's assets at baseline and the distance to the closest formal care provider; indicators for the child's age, relation to the household head, and gender; and the gender, age, and literacy level of the household head.[24] Lastly, specifications (4) and (6) include interactions of the treatment dummies with the "early" indicator, in order to assess if the policy interventions affect care seeking differently on early and care-required days. We re-estimate the model under inclusion of left-censored spells (treating the first observed day as the first spell day) and find that the results are robust, with only small changes in the estimated coefficient sizes (see Appendix C, table 12).

Figure 6.1, in addition graphs the predicted care seeking probabilities on each spell day (using the coefficient estimates and the baseline hazards $\lambda_0(t)$), split into the groups that do and do not receive free care and into early and care-required days.

While the proportional hazard model is easy to interpret and estimate, it also imposes a very specific structure on the treatment effects. We provide some evidence that the proportionality assumption is a reasonable approximation in Appendix C, figure C.1. In addition, we estimate the linear non-proportional model, where we allow

[24]Of these controls, only assets and child age affect healthcare demand positively and significantly, while the distance to the closest provider as a measure of non-monetary costs has the expected (negative) effect, but is not significant.

Table 6: Estimates using Cox proportional hazard model, uncensored and right-censored spells.

COX HAZARD MODEL	(1)	(2)	(3)	(4)	(5)	(6)
Early: care not yet required (IMCI)			0.147***	0.150***	0.140***	0.159***
			(0.0200)	(0.0524)	(0.0207)	(0.0565)
Healthworker visits (HW)	0.868	0.725	0.769	0.774	0.628*	0.636*
	(0.208)	(0.182)	(0.182)	(0.196)	(0.157)	(0.167)
Early x HW (EHW)				0.960		0.958
				(0.491)		(0.537)
Free care (FC)	2.894***	2.707***	2.792***	2.892***	2.522***	2.644***
	(0.517)	(0.495)	(0.498)	(0.569)	(0.464)	(0.542)
Early x FC (EFC)				0.805		0.758
				(0.333)		(0.323)
HW and FC (HWFC)	1.034	1.304	1.147	1.068	1.525	1.453
	(0.287)	(0.380)	(0.320)	(0.320)	(0.452)	(0.457)
Early x HWFC (EHWFC)				1.528		1.303
				(0.920)		(0.859)
Total assets in US$ (log)		1.050**			1.048*	1.048*
		(0.0232)			(0.0229)	(0.0229)
Dist. to closest formal (log)		0.892			0.851	0.852
		(0.0981)			(0.0934)	(0.0935)
Child controls	-	YES	-	-	YES	YES
Household head controls	-	YES	-	-	YES	YES
Observations	20,599	18,080	20,599	20,599	18,080	18,080
Partial LL	-3022	-2608	-2880	-2879	-2479	-2479
Wald chi-square	61.80	87.49	259	267.4	257.5	270.5
p-value: HWFC+HW=0					0.793	0.660
p-value: HWFC+EHWFC+HW+EHW=0						0.649

Significance levels: *** p<0.01, ** p<0.05, * p<0.1.
Standard errors clustered at the compound level (in parentheses).

the treatment effects to vary by day and by IMCI "early" status for each spell day up to day 7, with later days grouped.[25] Figure 6.2 graphs the coefficient estimates and 95% confidence intervals. The first panel show the baseline day fixed effects, or the average probability of seeking care on each spell day in the control group, in the left graph for days classified as too early for care by the IMCI, and in the right graph for care-required days. The following three panels show the coefficients on the day-specific free-care, healthworker, and combined indicator variables, reflecting the change in the probability of receiving care due to these treatments. Again, the results are split into early and care required days. Note that the total probability of care seeking is the sum of the individual effects across panels. For example, the probability of seeking care on the first spell day in the free-care group when care is required is the sum of the day fixed effect and the Day X free care coefficient (top two panels, right). It is equal to about 13%. This is similar to the estimate from the proportional hazard model, as seen in the right panel of figure 6.1.

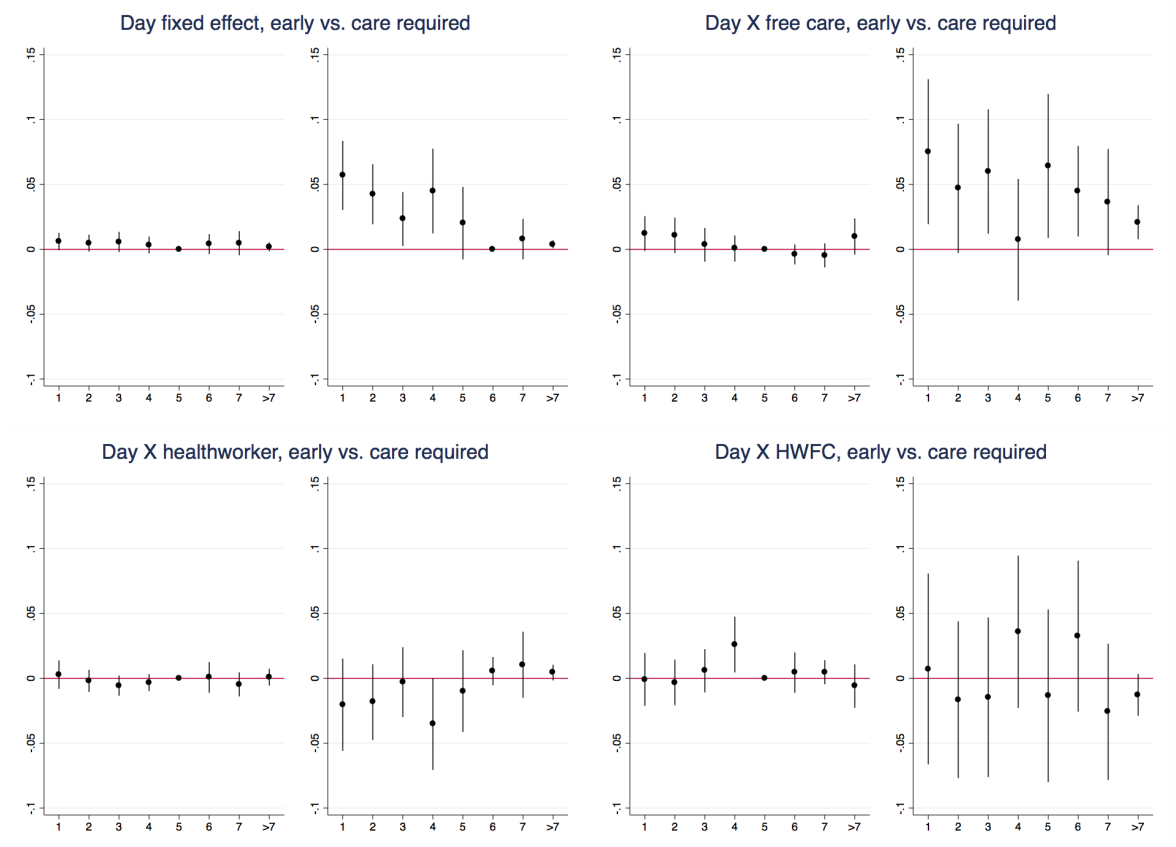
6.1 Results

Result 1: Overuse and Underuse in the Control Group. We find that demand for healthcare in the control group is low on care-required days and even lower on early days. The first panel of the non-proportional hazard estimates in figure 6.2 as well as the left panel of figure 6.1 show that the probability of seeking care on an early day is essentially zero. The probability on care-required days, when all children should seek care according to the IMCI, peaks at about 6% in day 1. Correspondingly, the highly significant coefficient on the "early" indicator variable in specifications (4) and (6) of the Cox model in table 6 shows that the hazard of care on early days is less than 15% of that on care-required days. These results show first that parents can distinguish early days from care-required days, and second, they behave in line with IMCI guidelines on early but not on care-required days – there is underuse but essentially no overuse. This suggests that parents in the control group apply a significantly higher cost threshold than the policymaker.

Result 2: The Effect of Free Care. The proportional hazard model shows that the probability of care-seeking nearly triples under free care. This can be seen from the coefficient estimates in specifications (1) and (2) of table 6. This increase leads

[25]Estimation results from a logit or nonlinear specification are very similar (not shown).

Figure 6.2: Coefficient plots from the non-proportional linear hazard model. Panel 1 shows the estimated probability of care seeking per spell day on early and on care-required days. Panels 2 and 3 show the estimated increase in the probability of care seeking on each spell day due to the free care and healthworker policies, and Panel 4 shows the additional effect in the group that receives both (each time split by early vs. care-required classification).

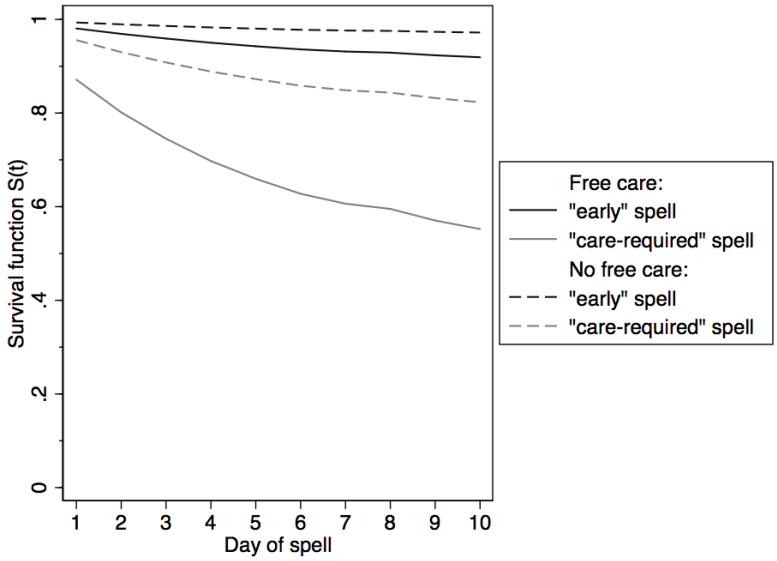


overwhelmingly to a reduction in underuse and little increase in overuse. From specifications (4) and (6), the hazard ratio for the interaction of "free care x early" is below one (though not significantly so). However, we have seen that without free care, care seeking occurs overwhelmingly on care-required days. This means that, if the *proportional* impact of free care is roughly the same on early and care-required days, the *absolute* impact of free care is much larger on care-required days.Figure 6.1 shows that free care increases the probability of seeking care on care-required days from about 4% to about 13% on day 1, and about 1% to 5% on day 6. The absolute impact on 'early' days is much smaller, with an increase from about 0.25% to roughly 1% on day 1, and from around 0.1% to 0.25% on day 6. The estimates from the non-proportional hazard model, as shown in the second panel of figure 6.2 confirms that free care has a small and insignificant estimated effect on the probability of care seeking on early days, while it has a much larger and (generally) significant effect on care required days.

To understand the implications of these estimates, we can simulate the overall probability of seeking care for different types of (hypothetical) illness spells, for example, an illness spell that immediately enters care-required status, compared to an illness spell that is 'early' throughout. These would correspond to a malaria spell versus a cold or cough, which may carry on for up to two weeks without requiring medical attention. Figure 6.3 depicts the simulated survival (i.e. the probability that care has not been sought by day t) for these two types of spell with and without free care up to day 10. The graph shows that the chance of a child getting care when not required remains below ten percent, even with free care. At the same time, the probability of receiving care in the severe spell rises to nearly 40% in the free-care group, relative to less than 20% in the group without free care.

The hazard model results suggest that by far the biggest impact of free care is to reduce underuse, with little increase in overuse. As yet another check, table 7 reports survival for the spells actually observed in our data. It shows the percentage of spells that end on a care-required vs. an early day, with and without a consultation, in the two treatment groups. The probability of receiving care conditional on a spell entering "care-required" status rises from 11% to 29% with free care, a substantial, but incomplete, reduction in underuse. Among the spells that end on an "early" day, the proportion receiving care only rises from 2.5% to 7.1%. From a policy perspective, these numbers are remarkable. Relative to the IMCI guidelines, free care reduces underuse in spells in which care is required by 20 percentage points, but increases the occurrence

Figure 6.3: Simulated survival probabilities for an "early" and a "care required" spell, with and without free care (for a child with the median sample characteristics and 50% probability of being in the healthworker treatment group, based on specification (3) in table 6).



of overuse by less than 5 percentage points.

Among spells that end in a consultation, the proportion that are in early status, and would therefore be classified as overuse, is virtually identical in the free care and no free care groups, at about 17%. On the one hand, this is a positive finding, because it means that there is no *relative* increase in wasteful spending (and as discussed earlier, our method of classification means that this proportion is an upper bound on overuse). An implication is that expanding free care to the other treatment groups would have lead to 191 additional doctor visits, of which at least 82% (157 visits) would have occurred when care is required. On the other hand, this proportion actually appears high, given that demand for care on any given early day is close to zero, and only 3.2% of all spells end in overuse. A reason for this disparity is the relatively high incidence of *underuse* among the care-required spells. Suppose for instance that the proportion of care-required spells that receive care goes from 29.3% up to 92.9%, for equal rates of underuse and overuse. Then the proportion of consultations that are appropriately targeted would be 94%. To achieve the same targeting rate by reducing overuse further, the proportion of early spells with consultations would have to be less than 2.4%.

Result 3: The Effect of Healthworkers. Table 6 shows that better information

Table 7: Spells by treatment group, percent ending on a care-required or an early day, and with or without consultation (uncensored vs. censored).

All spells: N = 3167		Unconditional		Conditional on	
	On last spell day:	Care Required	Early	Care Required	Early
Unconditional:	No Consultation	44.0%	42.8%	80.0%	95.2%
	Consultation	11.0%	2.2%	20.0%	4.8%
Conditional on a consultation:		83.5%	16.5%		
No free care: N = 1578		Unconditional		Conditional on	
	On last spell day:	Care Required	Early	Care Required	Early
Unconditional:	No Consultation	49.5%	43.4%	89.3%	97.4%
	Consultation	6.0%	1.1%	10.7%	2.6%
Conditional on a consultation:		83.9%	16.1%		
Free care: N = 1589		Unconditional		Conditional on	
	On last spell day:	Care Required	Early	Care Required	Early
Unconditional:	No Consultation	38.6%	42.2%	70.7%	92.9%
	Consultation	16.0%	3.2%	29.3%	7.1%
Conditional on a consultation:		83.3%	16.7%		

does not increase, and may reduce, demand for formal care. In all specifications, the coefficient on the healthworker dummy is negative, and significantly so in specifications (5) and (6). We argued in section 3.3 that the healthworkers may reduce overuse and underuse if cost thresholds between policymaker and parents are aligned but parents are poorly informed. However, the Cox estimates suggest that the healthworkers have proportionally the same effect on care seeking on early and care required days, and in absolute terms this means again that health workers reduce care seeking *by more* on care-required days. Panel 3 of figure 6.2 correspondingly shows negative point estimates for the healthworker effects on care-required days (some of which are significant at the 10% level), with (a tightly-estimated) zero effect on early days.

Result 4: The Interaction of Free Care and Healthworkers. Neither the Cox model nor the non-proportional hazard model suggest a strong interaction effect between healthworkers and free care. However, the point estimates from specifications (5) and (6) in table 6 imply that the negative healthworker effect is reversed in the group that receives free care as well. The hazard ratio on the healthworker effect in the free care group is 0.96 in (5) and 0.92 vs. 1.15 on care-required vs. early days, respectively, in (6). None of these are significantly different from 1, as shown by the p-tests in the last row of the table. This suggests that within the free-care group, utilization is the same among families who do and do not get healthworker visits.

These results are consistent with the view that the primary barrier to effective care seeking in our study population is costs, not information. Care seeking is much

Table 8: Pre-care spells by treatment group.

Pre-care spells	Control	Health worker	Free care	HW & FC	Total
Uncensored and right-censored spells					
Mean length	7.014	6.959	5.810***	6.402*	6.543
SD	(7.341)	(7.924)	(6.027)	(7.103)	(7.142)
N	781	790	798	791	3160
Within spell: early days - care not yet required (IMCI)					
Mean	3.643	3.037***	3.038***	3.154**	3.216
SD	(4.304)	(3.939)	(3.774)	(4.159)	(4.053)
Within spell: care-required days (IMCI)					
Mean	3.371	3.923	2.772**	3.248	3.327
SD	(6.328)	(7.358)	(5.119)	(6.207)	(6.311)
Uncensored spells only					
Mean length	3.764	6.346*	4.401	4.959	4.762
SD	(4.303)	(8.733)	(4.861)	(7.100)	(6.278)
N	55	52	157	148	412

*Significance levels: *** 1%, ** 5%, * 10%, t-test on mean difference from control.*

less likely on 'early' days compared to 'care required' days, and free care does not lead to large increases in overuse, suggesting that parents can identify when there is no need for care yet. However, parents initially have a cost threshold far above that of the policymaker, and even when care is subsidized, alignment is not complete, so that a significant amount of underuse remains. Likely reasons are the significant non-pecuniary costs of care – for example the opportunity cost of missed work, and the effort involved in transporting the child to the clinic – as well as the remaining monetary costs resulting from unsubsidized care components (see section 7).

There is little evidence that healthworkers can reduce the misallocation of health resources, as might be expected in a situation in which parents are relatively well-informed to begin with. Parents in the sample are rarely seeking care too early, so there is very little overuse for the healthworker to correct. Healthworkers appear in fact to *increase* underuse when there is no subsidy. We have shown that this can occur if the cost thresholds of parents and policymaker are not aligned, and the parents use the information from the healthworkers to learn when an illness is *not* extremely serious, leading them to use less healthcare (see section 3.3).[26]

6.2 The Magnitude of Censoring and Selection Effects and their Role for Estimating Demand

In section 5 we discussed the potential biases that can arise from censoring and selection. Here we demonstrate that one would draw radically different policy lessons if these issues were ignored.

The top rows of table 8 report average observed lengths for the spells used in the hazard analysis. As argued earlier, average spell length underestimates time-to-care due to censoring, and the censoring bias is larger when true time-to-care is on average higher. As an illustrating example, suppose the control group seeks care on average after ten days of illness, and the treatment group after six days, but the majority of untreated spells end after seven days or earlier. In this case a lot of the "latent" increase in demand for care will go unobserved, and the control spells are censored much more often than the treatment spells. Furthermore, censoring will over-proportionally affect the 'care-required' portion of the spell, which always occurs at the end.

Table 8 shows that the effect of free care on average spell length in our data is indeed small, given that the probability of care seeking nearly triples according to the hazard estimates. The effect of free care on early days is also more robust than on care-required days. Naively interpreted, this would suggest that subsidies mainly lead to wasteful doctor visits, while care-seeking behavior on care-required days is unaffected. Yet we know that the *conditional* probability to seek care is nearly seven times higher on care-required than on early days; the difference arises because underuse is still high and most spells enter the "care required" stage not at all or only after several days.

We argued earlier that selection effects make it even more problematic to estimate time-to-care from uncensored spells only, i.e. spells that end in a consultation, for example from data collected at the provider, rather than the household. The bottom row of table 8 shows that this concern is of practical importance: the number of uncensored spells nearly triples in the two free-care groups, but the average time to care in those uncensored spells is actually *longer* than in the control. We can assess the magnitude of the bias by reestimating the hazard model on the uncensored spell data only. The results, shown in table 13 in Appendix C, indicate that we would have

[26] This interpretation is also consistent with a scenario in which the parents do not learn from the healthworker how to interpret symptoms themselves, but use the signal from the regular healthworker's assessment of the child to assess illness severity. The healthworker visit is essentially used as a substitute for a formal consultation.

wrongly concluded that free care has no or even a negative effect on time-to-care, and that healthworker visits alone lead to further substantial delay. Using only uncensored data also almost doubles the estimated hazard ratio on early days, so that parents appear less able to discern the health care needs of their children. This bias again arises because changes to care-seeking on early days are more frequently observed than on care-required days.

These findings make an important methodological point: in order to estimate the unbiased effect of a policy intervention on health care demand conditional on need for care, it is imperative to collect data at the household level and estimate demand with a hazard model. In a population like ours where underuse is rampant, data collected at the point of use (e.g. the healthcare provider or insurer) is subject to selection effects that are so strong that they reverse the estimated impact of free care on care-seeking behavior. Another important reason to estimate demand conditional on health status is the ability to conduct counterfactual policy analysis and out-of-sample predictions. We can, for example, predict healthcare use for populations with different levels of morbidity, and for specific illnesses with known expected symptom paths. An example was shown in figure 6.3 above. It is easy to construct similar paths for more complicated diseases, possibly using more precise coefficient estimates of care seeking for different symptom histories. We believe this to be a promising avenue for future research. The results could be immensely useful for healthcare policy as well as public health – as one example, a more detailed version of our model could be used to make predictions about the care-seeking behavior of Ebola patients and thereby inform epidemiological research on the spread of new diseases or drug-resistant parasites.

7 Demand Effects on Costs and Health Outcomes

In order to evaluate the impact of the demand changes induced by Action for Health, we present some results on its impact on cost and health outcomes.

7.1 Private and Social Healthcare Spending

Table 9 reports the effect of the program on costs. Each column represents an OLS regression with a different measure of cost as the dependent variable and the treatment dummies as the independent variables (with standard errors clustered at the compound

Table 9: Regressions of cost of care and value of care onto treatment indicators.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Private cost					Social cost		
	per visit	per child				per visit	per child	
	CSCOM	Total	CSCOM	Other formal	Informal	CSCOM	Total	CSCOM
Healthworker	642.0	444.4	137.2	219.0	4.073	-62.73	326.9	51.91
visits (HW)	(585.9)	(270.5)	(137.8)	(178.9)	(67.09)	(550.6)	(290.2)	(156.0)
Free care (FC)	-1,734***	-16.95	176.5	-121.0	-73.08	96.07	1,468***	1,666***
	(429.9)	(253.6)	(140.7)	(168.3)	(68.93)	(398.9)	(373.6)	(279.3)
HW and FC	-753.1	-622.9*	-218.9	-303.4	-2.278	-122.5	-506.3	-165.0
(HWFC)	(641.1)	(364.1)	(210.2)	(222.9)	(100.8)	(633.4)	(504.6)	(375.5)
Constant	2,681***	1,296***	355.8***	470.8***	358.6***	3,946***	1,524***	523.6***
	(385.3)	(170.0)	(71.37)	(132.6)	(48.73)	(328.0)	(185.7)	(86.03)
Observations	518	1,552	1,552	1,552	1,552	518	1,552	1,552
R-squared	0.109	0.005	0.001	0.006	0.002	0.001	0.022	0.062

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Private costs for consultation and treatment as reported by the parents.

Social costs imputed from private cost and median prices by consultation type and treatment received.

level). The first column reports private costs (paid by the household) for an individual CSCOM visit. The next four columns report the effects on private spending per child, first in total and then for CSCOM, other formal, and informal care separately. The next columns first report the social costs per CSCOM visit, and then total and CSCOM social costs per child. The social cost is given by the undiscounted price of care (which consists of all private costs born by the parents or other contributors, e.g. relatives, plus the costs incurred by Mali Health) and provides a proxy for the value of healthcare consumed.[27]

We use these results to answer three questions about the impact of the program on costs.

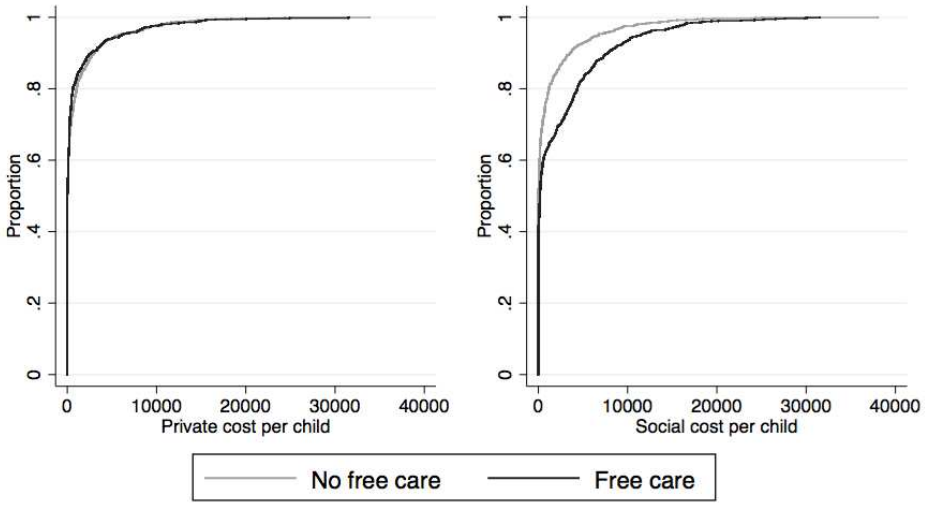
[27]The source for all cost measures are the recorded payments made by the family for consultations and medications. Social costs are calculated using prices recorded for those purchases where the respondent reported having paid the full price themselves. We use these prices combined with medication brand names and point of purchase to construct the median price of each medication by source. In addition, we use provider information to construct the median price of a consultation by source. If the respondent reported another person paying for care or received CSCOM care in the free care group, we compare the sum of median prices for the services received with the private expenses the respondent reported. If the latter is higher, the social cost is assumed to be equal to the reported private cost, if it is lower, we assume the social cost to be the imputed median price. The amount calculated in this manner is intended to reflect the social cost of care, regardless if it was paid by the family, the NGO, or a third party. – A potential worry for the assessment of value is that the program may lead to price changes, or substitution between vendors who supply the same treatment at different prices. Furthermore it is unclear how informal care should be valued. However, pricing treatment identically across vendors or excluding informal care does not substantively change the results (analysis not shown).

Question 1: Does Free Care Crowd Out Private Spending or Increase Care Consumed? The impact of free care on the total amount of healthcare received by the child depends on the fungibility of the subsidy. If families drastically reduce their health spending as a result of the program, then the total value of care received by the child may not change substantially. However, column (2) shows that this is not the case. The impact of free care on total private health expenditure is small in magnitude and insignificant. If anything, columns (3)-(5) suggest that there is somewhat *higher* expenditure per child at the CSCOM, only partially offset by lower spending for other formal and informal care. Column (1) shows that free care reduces the cost per CSCOM visit by about 1700 CFA, leaving a residual cost of approximately 950 CFA for care not covered by Action for Health. The increase from 0.13 to 0.55 visits per child combined with this residual means that private expenditure under the subsidy remains the same.

As a result, the value of healthcare received by children in the free care group is much higher. Column (7) of table 9 shows that the social cost per child roughly doubles under free care, from 1500 to 3000 CFA. Figure 7.1 further illustrates this point. The two panels show the CDFs of private and social costs per child for those with and without free care. The distribution of private costs is virtually unchanged, while the social cost distribution under free care first order stochastically dominates that without free care. The graphs also show that almost 50% of children do not consume any care, with or without the subsidy, because demand for care remains latent when the child is healthy. Conditional on experiencing a spell of illness the increase in the value of care consumed is therefore much greater.

Question 2: Do Healthworkers Affect Healthcare Expenditure? We have already seen that healthworkers do little to encourage formal healthcare seeking. Using table 9 we can ask if they have an impact on healthcare expenditure more generally. The point estimates for the effect of health workers on private expenditure are positive in all categories, but significant in none. Column (1) shows that the healthworkers cause an increase in cost per visit (we saw in the previous section that, if anything, healthworkers reduce the propensity of families to seek formal care). These spending increases occur only in the group that does not receive free care: The coefficient on the HWFC dummy is negative and of similar size to that on the HW dummy (and significant in the case of total expenditure).

Figure 7.1: The cumulative distribution function of private and social healthcare costs.



Question 3: Do the CSCOM Clinics Oversell to the Free Care Group? In general, one might expect that providers respond to the financial incentives of health subsidies by raising treatment costs, but at least in the Action for Health program there is no such effect. Costs per visit are essentially unaffected at about CFA 4000. In Appendix C, Table 14 we show that this holds also when controlling for spell day and IMCI classification, and is thus not a consequence of selection effects. By extension, there is no obvious problem of moral hazard on the part of the provider. This is likely due to monitoring by Mali Health, and it allows us to study demand for healthcare holding the provider side fixed. The estimates in the appendix do reveal that CSCOM doctors prescribe treatment costing about CFA 1800 less to children who come visit on an early day, compared to when care is required. This lends additional credibility to the WHO guidelines as an indicator of need for care.

7.2 Health Outcomes

So far we have shown that free care dramatically increases the value of care consumed, but not how this affects health. Access to health care in our setting encompasses both evaluation by a doctor, and, conditional on evaluation, treatment. Treatment can affect health outcomes in some cases immediately, by leading to faster recovery and a reduction in symptoms. An important benefit of timely formal care, however, is to detect those (low-frequency) cases where treatment can prevent serious long-term

harm or death, as with malaria. Our study was not powered to detect the unconditional impact of the program on most health outcomes. More severe outcomes, such as the death of a child, occur too rarely in our sample. Day to day illness incidence is driven in large part by infection rates and other exogenous events, and the majority of sick children see a doctor only towards the end of an illness spell or not at all. As a consequence, the effect of the program on average days of sickness is expected to be small. Indeed, the number of spells in the four groups is nearly identical (table 8), and the full program group actually has the *highest* average number of sickness days of the four treatment groups (though this is driven by a higher incidence of wounds, burns and injuries, see table 11 in the appendix).

Nonetheless, our data indicates some positive effects on average total illness spell length (that is, persistence of the illness conditional on a spell occurring) as well as the mother's concern about the child as a proxy for the (expected) utility experienced by the child's caregiver. Table 10 shows estimates of the effect of free care on the average duration of an illness spell (including days after care was sought) as well as on the daily probability of the mother being 'concerned' or 'very concerned' about her child's health (a variable collected as part of the health diaries). Regressions (3) and (4) include all days, (5) and (6) only days with symptoms. Each pair of regression equations first shows a simple OLS estimate, and then an OLS with household and child control variables. In (5) and (6), we can include an "early" indicator, and an interaction of free care with "early". All standard errors are clustered at the compound level.

The OLS estimates suggest that being in one of the free-care groups is associated with slightly shorter average spell length. Since only a small percentage of spells receive formal care, the effect of care on the individual spell is larger (but note that there is selection into care). The effect is consistent with findings from Tanzania, where the use of formal healthcare is associated with shorter illness duration and lower incidence of fever and malaria (Adhvaryu and Nyshadham [2015]). Free care also reduces the average proportion of illness days when the mother is concerned or very concerned about the child significantly from 30% to about 22%. The effect is smaller but still significant when including all observed days. Note also that mother concern is significantly lower on early days, again consistent with the idea that mothers agree with the WHO assessment about a child's health based on observed symptoms.

Table 10: Effect of free care on average illness spell length and mother concern.

	Spells		All days		Illness days only	
	Illness spell length		Mother concerned or very concerned		Mother concerned or very concerned	
	(1)	(2)	(3)	(4)	(5)	(6)
Free care (FC)	-0.169	-0.866*	-0.0186**	-0.0189**	-0.0827***	-0.0772*
	(0.407)	(0.513)	(0.00808)	(0.00889)	(0.0313)	(0.0453)
Early x FC (EFC)					0.0366	0.0322
					(0.0367)	(0.0390)
Early: care not yet required (IMCI)					-0.119***	-0.114***
					(0.0273)	(0.0285)
Constant	7.357***	6.313***	0.0734***	0.0461*	0.291***	0.300***
	(0.328)	(1.363)	(0.00663)	(0.0247)	(0.0252)	(0.104)
Controls		Yes		Yes		Yes
Observations	3,144	2,809	98,537	85,299	27,859	25,069
R-squared	0.000	0.007	0.001	0.006	0.022	0.029

*** p<0.01, ** p<0.05, * p<0.1

Standard errors clustered at the compound level (in parentheses). Controls include child and household characteristics.

8 Conclusion

The policy debate around the subsidization of healthcare has been subject to large swings. The Bamako consensus in the 1980s, which led to a broad move towards user-fee financed healthcare systems, is in the process of reversing: several aid organizations now advocate free care (UK Secretary of State for International Development [2009], McPake et al. [2008]), and a number of African countries have introduced (partially) free health care especially for mothers and small children (Ridde and Morestin [2012], Yates [2007]) in an effort to reduce high mortality and morbidity rates in these populations. However, as others have pointed out (Powell-Jackson et al. [2014]), there is relatively little convincing evidence of the effect of abolishing user fees on utilization or ultimately health outcomes.

This study aims to open the "black box" of healthcare demand and understand healthcare use conditional on health status and need for care according to medical standards. We model the dynamic problem that parents face when choosing to visit a doctor, and show that hazard of care analysis needs to account for spell censoring due to spontaneous recovery to avoid biased effect estimates. Conditioning demand estimates on the presence of negative health shocks as the drivers of healthcare demand may resolve the puzzle of low uptake and low utilization effects of reducing healthcare costs in previous health insurance trials, and improve their external validity and precision.

Our results have encouraging implications for the user-fee debate. The welfare cost of subsidies could be prohibitive if they lead to substantial mis-targeting and overuse of healthcare, risking to overburden the severely limited capacity of healthcare systems in poor countries and thus diverting resources away from those truly in need. However, we find that healthcare usage nearly triples in response to subsidies, but does not result in substantial overuse, because demand comes largely from children for whom medical care is required according to WHO guidelines. At the same time, the families in our study are quite adept at recognizing their child's need for healthcare, so that additional information policies do not markedly improve the allocation of care. Overall, it appears that overuse and moral hazard or insufficient health knowledge are not a primary concern in this population, whereas underuse due to diverging spending priorities between parents and policymaker remains high. Thus, policy efforts should be focused on increasing utilization when care is required, rather than further curbing the existing low levels of overuse.

References

Achyuta Adhvaryu and Anant Nyshadham. Return to treatment in the formal health care sector: Evidence from Tanzania. *American Economic Journal: Economic Policy*, 7(3):29–57, 2015.

John Akin, Nancy Birdsall, and David de Ferranti. *Financing Health Services in Developing Countries: An Agenda for Reform*, volume 34. World Bank, 1987.

John S. Akin, Charles C. Griffin, David K. Guilkey, and Barry M. Popkin. The demand for primary health care services in the Bicol region of the Philippines. *Economic Development and Cultural Change*, 34(4):755–782, 1986.

Aviva Aron-Dine, Liran Einav, Amy Finkelstein, and Mark Cullen. Moral hazard in health insurance: Do dynamic incentives matter? *Review of Economics and Statistics*, 97(4):725–741, 2015.

Nava Ashraf, James Berry, and Jesse M. Shapiro. Can higher prices stimulate product use? Evidence from a field experiment in Zambia. *American Economic Review*, 100 (5):2383–2413, 2010.

Nava Ashraf, B. Kelsey Jack, and Emir Kamenica. Information and subsidies: Complements or substitutes? *Journal of Economic Behavior and Organization*, 88:133–139, 2013.

Zulfiqar A Bhutta, Samana Ali, Simon Cousens, Talaha M Ali, Batool Azra Haider, Arjumand Rizvi, Pius Okong, Shereen Z Bhutta, and Robert E Black. Interventions to address maternal, newborn, and child survival: what difference can integrated primary health care strategies make? *The Lancet*, 372(9642):972–989, 2008.

Marika Cabral. Claim timing and ex post adverse selection: Evidence from dental "insurance". *mimeo*, 2012.

Jessica Cohen, Pascaline Dupas, and Simone Schaner. Price subsidies, diagnostic tests, and targeting of malaria treatment: evidence from a randomized controlled trial. *American Economic Review*, 105(2):609–645, 2015.

Janet Currie, Wanchuan Lin, and Wei Zhang. Patient knowledge and antibiotic abuse: evidence from an audit study in China. *Journal of Health Economics*, 30:933–949, 2011.

Jishnu Das. The quality of medical care in low-income countries: From providers to markets. *PLoS Medicine*, 8(4):1–2, 2011.

Jishnu Das and Jeffrey Hammer. Which doctor? combining vignettes and item response

to measure clinical competence. *Journal of Development Economics*, 78(2):348–383, 2005.

Jishnu Das and Jeffrey Hammer. Location, location, location: residence, wealth, and the quality of medical care in Delhi, India. *Health Affairs*, 26(3):w338–w351, 2007.

Jishnu Das and Jeffrey Hammer. Quality of primary care in low-income countries: Facts and economics. *Annual Review of Economics*, 6:525–553, 2014.

Jishnu Das and Thomas Pave Sohnesen. *Patient Satisfaction, Doctor Effort and Interview Location: Evidence from Paraguay*. World Bank, 2006.

Jishnu Das, Jeffrey Hammer, and Carolina Sánchez-Peramo. The impact of recall periods on reported morbidity and health seeking behavior. *World Bank Policy Research Working Paper*, 5778, 2011.

Jishnu Das, Alaka Holla, Aakash Mohpal, and Karthik Muralidharan. Quality and accountability in healthcare delivery: Audit-study evidence from primary care in India. *NBER Working Paper*, 21405, 2015.

W. Dow, P. Gertler, R.F. Schoeni, J. Strauss, and D. Thomas. Health care prices, health and labor outcomes: experimental evidence. *RAND Working Paper*, 97-01, 2001.

Pascaline Dupas. Short-run subsidies and long-run adoption of new health products: Evidence from a field experiment. *Econometrica*, 82(1):197–228, 2014.

Pascaline Dupas and Jessica Cohen. Free distribution or cost-sharing? Evidence form a randomized malaria prevention experiment. *Quarterly Journal of Economics*, 125 (1):1–45, February 2010.

Amy Finkelstein, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Heidi Allen, Katherine Baicker, and Oregon Health Study Group. The Oregon health insurance experiment: Evidence from the first year. *Quarterly Journal of Economics*, 127(3):1057–1106, 2012.

Greg Fischer, Dean Karlan, Margaret McConnell, and Pia Raffler. To charge or not to charge: evidence from a health products experiment in Uganda. *NBER Working Paper*, 20170, 2014.

Andrew D. Foster. Prices, credit markets and child growth in low-income rural areas. *Economic Journal*, 105(430):551–570, 1995.

P. Gertler and J. Van der Gaag. *The willingness to pay for medical care: evidence from two developing countries*. Baltimore and London, 1990.

Paul Gertler, Luis Locay, and Warren Sanderson. Are user fees regressive? The welfare

implications of health care financing proposals in Peru. *Journal of Econometrics*, 36 (1-2):67–88, 1987.

Donna B. Gilleskie. A dynamic stochastic model of medical care use and work absence. *Econometrica*, 66(1):1–45, 1998.

Mark C. Hornbeck, Arnold V. Hurtado, and Richard E. Johnson. Health care episodes: definition, measurement and use. *Medical Care Research and Review*, 42(2):163–218, 1985.

Emmanuel Jimenez. *Human and physical infrastructure: public investment and pricing policies in developing countries*, volume 3, chapter 43, pages 2773–2843. Elsevier, 1995.

Andrew M. Jones. Health econometrics. In A. J. Culyer and Joseph P. Newhouse, editors, *Handbook of Health Economics*, volume 1, chapter 6, pages 267–344. Elsevier, 2000.

Emmett B. Keeler, Joan L. Buchanan, John E. Rolph, Janet M. Hanley, and David M. Reboussin. The demand for episodes of medical treatment in the health insurance experiment. *RAND Working Paper*, 1988.

Gary King, Emmanuela Gakidou, Kosuke Imai, Jason Lakin, Ryan T. Moore, Clayton Nall, Nirmala Ravishankar, Manett Vargas, Martha M. Téllez-Rojo, Juan E. Hernández Ávila, Mauricio Hernández Ávila, and Hector Hernández Llamas. Public policy for the poor? a randomised assessment of the Mexican universal health insurance programme. *The Lancet*, 373(9673):1447–1454, 2009.

Amanda Kowalski. Censored quantile instrumental variable estimates of the price elasticity of expenditure on medical care. *Journal of Business and Economic Statistics*, forthcoming.

Michael Kremer and Rachel Glennerster. Improving health in developing countries: evidence from randomized evaluations. In *Handbook of Health Economics*, volume 2, chapter 4, pages 201–315. Elsevier, 2012.

Kenneth L. Leonard and Melkiory C. Masatu. Using the Hawthorne effect to examine the gap between a doctor's best possible practice and actual performance. *Journal of Development Economics*, 93:226–234, 2010.

K.L. Leonard, G.R. Mliga, and D. Haile Mariam. Bypassing health centres in tanzania: revealed preferences for quality. *Journal of African Economies*, 11(4):441, 2002.

Jenny I. Litvack and Claude Bodart. User fees plus quality equals improved access to health care: Results of a field experiment in cameroon. *Social Science and Medicine*, 37(3):369 – 383, 1993.

Willard G. Manning, Joseph P. Newhouse, Naihua Duan, Emmett B. Keeler, and Arleen Leibowitz. Health insurance and the demand for medical care: Evidence from a randomized experiment. *American Economic Review*, 77(3):251–277, 1987.

Diane McIntyre, Michael Thiede, Göran Dahlgren, and Margaret Whitehead. What are the economic consequences for households of illness and of paying for health care in low- and middle-income country contexts? *Social Science and Medicine*, 62(4): 858–865, 2006.

Barbara McPake, Alice Schmidt, Edson Araujo, and Christine Kirunga-Tashobya. *Freeing up Healthcare: A guide to removing user fees*. Save the Children UK, London, 2008.

E. Miguel and M. Kremer. Worms: identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, 72(1):159–217, 2004.

J. Nabyonga, M. Desmet, H. Karamagi, P. Y. Kadama, F. G. Omaswa, and O. Walker. Abolition of cost-sharing is pro-poor: evidence from Uganda. *Health Policy and Planning*, 20(2):100–108, 2005.

Hoa L. Nguyen, Jane S. Saczynski, Joel M. Gore, and Robert J. Goldberg. Age and sex differences in duration of prehospital delay in patients with acute myocardial infarction: A systematic review (pmc3072277). *Circulation: Cardiovascular Quality and Outcomes*, 3:82–92, 2010.

Timothy Powell-Jackson, Kara Hanson, Christopher J.M. Whitty, and Evelyn K. Ansah. Who benefits from free healthcare? evidence from a randomized experiment in Ghana. *Journal of Development Economics*, 107:305–319, 2014.

Valéry Ridde and Florence Morestin. A scoping review of the literature on the abolition of user fees in health care services in Africa. *Health Policy and Planning*, 27(3):1–11, 2012.

Alfonso Rosales and Kristin Weinhauer. *C-IMCI: Community Integrated Management of Childhood Illness*. Catholic Relief Services, 2003.

David E. Sahn, Stephen D. Younger, and Garance Genicot. The demand for health care services in rural Tanzania. *Oxford Bulletin of Economics and Statistics*, 65(2): 241–260, 2003.

João Santos Silva and Frank Windmeijer. Two-part multiple spell models for health care demand. *Journal of Econometrics*, 104:67–89, 2001.

G. L. Stoddart and M. L. Barer. Analysis of demand and utilization through episodes of medical services. *McMaster University working paper*, 81-06, 1981.

Dag Gundersen Storla, Solomon Yimer, and Gunnar Aksel Bjune. A systematic review of delay in the diagnosis and treatment of tuberculosis. *BMC Public Health*, 8(15), 2008.

Rebecca L. Thornton, Laurel E. Hatt, Erica M. Field, Mursaleena Islam, Freddy Solís Diaz, and Martha Azucena González. Social security health insurance for the informal sector in Nicaragua: a randomized evaluation. *Health Economics*, 19(S1): 181–206, 2010.

UK Secretary of State for International Development. *Eliminating World Poverty: Building our Common Future*. Department for International Development White Paper, London, 2009.

United Nations Inter-Agency and Expert Group on MDG Indicators. *The Millennium Development Goals Report 2015*. United Nations, 2015.

WHO. *Integrated Management of Childhood Illness: Chart Booklet*. World Health Organization, 2014.

WHO Department of Child and Adolescent Health and Development. *Handbook IMCI: Integrated Management of Childhood Illness*. World Health Organization and Unicef, 2005.

WHO Global Malaria Programme. *World Malaria Report 2014*. World Health Organization, 2015.

Rob Yates. The impact of abolishing user fees in Africa – recent developments in six African countries. *SSRN Working Paper*, June 2007.

A Examples and Proofs

Proof of Lemma 1

Note that Bayesian updating implies

$$P_{t+1}(I_n) = \frac{\pi(I_n)P_t(I_n)}{\sum_{m=1}^{N}\pi(I_m)P_t(I_m)} = \frac{\pi(I_n)P_t(I_n)}{\pi_t(P_t)}.$$

Thus, the probability of illness I_n (strictly) decreases if its recovery rate is above average, and increases if it is below average. This means that the time-t distribution of recovery rates first-order stochastically dominates the time $t+1$ distribution, and $\pi_t(P_t)$ must increase over time.

Proof of Proposition 1

We show that the parent has no desire to deviate from this strategy given that it is used in all future periods. First consider the choice of the parent when beliefs first cross the threshold, so that $\pi_t(P_t) < K$ but $\pi_{t+1}(P_{t+1}) \geq K$. In this case, they can either choose to go to the doctor in the current period, and pay cost $-C$, or wait and follow the equilibrium strategy to go in the next period if the child is still sick. The latter has an expected cost of $\pi_t(P_t)\delta(-S-C)$, so $\pi_t(P_t) < K$ ensures that it is optimal not to consult a doctor today. Since $\pi_{t+1}(P_{t+1}) \geq K$ and π_t weakly increases over time, the same logic ensures that it is also optimal to go immediately in $t+1$ and any period after. Now consider $t-1$. Here, the choice is between going immediately or waiting for two periods before receiving treatment. The utility from waiting is

$$\begin{aligned}\pi(P_{t-1})\delta(-S+\pi(P_t)\delta(-S-C)) &< \\ \pi(P_{t-1})\delta(-S-C) &\leq -C\end{aligned}$$

where the first inequality uses that $\pi(P_t)\pi_t$ is below the threshold K. Iterating this argument shows that it is optimal to wait in all earlier periods $t-k < t-1$ where $\pi_{t-k}(P_{t-k})$ is below K.

Details for Examples 1 and 2

In example 1, I_S and I_L are equally likely, so that $\pi_t = \frac{1}{2}(\pi(I_S) + \pi(I_L))$. Letting $\epsilon > 0$ such that $\pi(I_S) = \pi_t + \epsilon$ and $\pi(I_L) = \pi_t - \epsilon$, Bayesian updating implies that $\pi_t < \pi_{t+1} = \frac{\frac{1}{2}\pi(I_S)^2 + \frac{1}{2}\pi(I_L)^2}{\pi_t} = \pi_t + \frac{\epsilon^2}{\pi_t}$.

In example 2, we have $\pi_t^{*L} = \pi(P_t(\gamma^L)) = \pi_t - \frac{1}{2}\epsilon < \pi_t$, and $\pi_{t+1}^{*L} = \pi(P_{t+1}(\gamma^L)) = \pi_{t+1} - \frac{\epsilon}{\pi_t} < \pi_t$: the probability of not recovering is increasing, but actually lower than the initial π_t in either period. Under γ^{*S}, the probabilities of illnesses I_S and I_L are reversed, so that $\pi_t^{*S} = \pi(P_t(\gamma^{*S})) = \pi_t + \frac{1}{2}\epsilon$ and $\pi_{t+1}^{*S} = \pi(P_{t+1}(\gamma^{*S})) = \pi_{t+1} + \frac{\epsilon}{\pi_t}$.

B Health Diary, Symptoms, and IMCI Classifications

Figure B.1 shows the health diary. The diary has entries for nine major symptoms, the mother's concern about the child, and doctor and pharmacy visits.

The full list of 12 symptoms were chosen based on the Integrated Management of Childhood Illness (IMCI) guidelines (see WHO Department of Child and Adolescent Health and Development [2005], WHO [2014]). IMCI is a joint program of the World Health Organization and Unicef and aims at establishing simple procedures for the management of the main causes of morbidity and mortality of small children. The guidelines differentiate between low and high malaria regions. Rosales and Weinhauer [2003] have adapted the IMCI for use by community healthworkers, who typically have no formal medical training, but can be instructed to follow simple protocols for the basic diagnosis of illness. These protocols are mainly designed to detect a need for formal medical evaluation. Symptoms are classified into acute danger signs and symptoms that point to a particular illness or class of illnesses, e.g. malaria or generalized fever. The choice of symptoms collected was based on ease of observation and description for both surveyors and mothers, and explicit rules found in the IMCI guidelines for referral to formal care. The symptoms are:

– Convulsions, fits, or spasms (danger sign)
– Lethargic or less conscious (danger sign)
– Unable to drink or breastfeed (danger sign)
– Vomiting everything (danger sign)
– Coughing (respiratory disease)
– Difficulty breathing (respiratory disease)

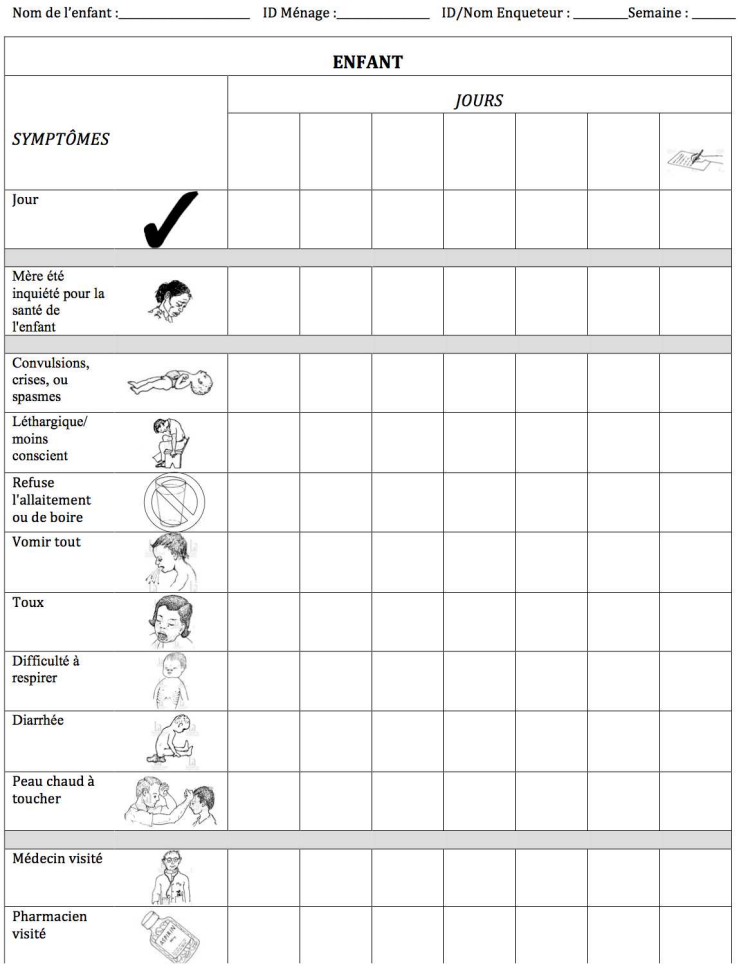


Figure B.1: Sample of the health diary. Along the top, the surveyors enter the days of the week. The last day of each calendar, marked by a "writing" symbol, corresponds with the visit of the surveyor, the first day corresponds with the previous visit and used to record events occurring after the surveyor leaves. Mothers were asked to use the diaries as a memory aid, but records were also taken if the mother did not fill in the health diary.

– Diarrhea – If diarrhea: more than three loose stools per day? (diarrheal disease)

– If diarrhea: blood in the stool? (indicator for dysentery)

– If diarrhea: sunken eyes? (indicator for dehydration)

– Unusually hot skin (under 2 months age: unusually cold skin) (fever)

Additionally, we manually classified symptoms recorded under "other health changes" into "cold symptoms", "ear pain", "skin rash", "wound or injury", "unusually hot" (typically recorded as "fever" or "malaria"), "head/neck/eye pain", "stomach pain", and "other".

Table 11 shows the frequency of reported symptoms by randomized treatment group and in total. The full treatment group experiences or reports more illness compared to the control, significant at the 10% level, although some symptoms are also reported less often, for example "sunken eyes". This difference may be a statistical accident, but may also be an effect of the treatment if both free care and healthworkers increase parents' attention to their child's symptoms. It would be hard to say if symptoms in the full treatment group are "over-reported" or in the control "underreported". Note, however, that the difference is driven by injuries and cold symptoms; Mali Health does not pay for healthcare related to either. Moreover, the incidence of illness *spells* is almost identical in all treatment groups (see table 8).

Based on the C-IMCI, we used guidelines on urgent or non-urgent referral to a clinic to determine when a child should seek formal care. The following rules were applied:

– Any of the danger signs require immediate (same day) care.

– Diarrhea with blood in the stool or sunken eyes require immediate referral on suspicion of dysentery or severe dehydration, respectively.

– Diarrhea without signs of dysentery or dehydration requires non-urgent referral after at most 5 days of continual illness.

– Fever with a rash and cough or cold symptoms require immediate referral on suspicion of measles.

– Fever without cough, cold symptoms, difficulty breathing, rash, or ear infection requires immediate referral on suspicion of malaria.

– Any other fever requires a non-urgent referral for generalized fever.

– A simple cough requires non-urgent care after 14 days on suspicion of tuberculosis.

– Cold symptoms and difficulty breathing require non-urgent care after 14 days on suspicion of a bacterial rather than viral infection.

– Ear pain should lead to non-urgent referral for acute or chronic ear infection.

Table 11: Symptom incidence by treatment group.

	Control	Health worker	Free care	HW & FC	Total
Symptom days as proportion of observed days per child.					
Any symptom	**0.285**	**0.297**	**0.292**	**0.320***	**0.298**
	(0.258)	(0.260)	(0.259)	(0.257)	(0.258)
Convulsions, fits, or spasms	0.000	0.002	0.000	0.001*	0.001
	(0.006)	(0.032)	(0.003)	(0.010)	(0.017)
Lethargic or unconscious	0.013	0.010	0.014	0.019	0.014
	(0.070)	(0.034)	(0.056)	(0.049)	(0.054)
Unable to drink or breastfeed	0.005	0.003	0.004	0.003	0.003
	(0.036)	(0.016)	(0.018)	(0.016)	(0.023)
Vomiting everything	0.011	0.012	0.010	0.013	0.011
	(0.030)	(0.034)	(0.022)	(0.029)	(0.029)
Coughing	0.108	0.119	0.099	0.122	0.112
	(0.168)	(0.185)	(0.164)	(0.169)	(0.172)
Difficulty breathing	0.015	0.019	0.013	0.018	0.016
	(0.051)	(0.070)	(0.042)	(0.046)	(0.053)
> 3 loose stools	0.024	0.020	0.025	0.020	0.022
	(0.065)	(0.055)	(0.071)	(0.051)	(0.061)
Blood in the stool	0.003	0.002	0.001	0.001	0.002
	(0.015)	(0.017)	(0.011)	(0.011)	(0.014)
Sunken eyes	0.008	0.005	0.007	0.004*	0.006
	(0.038)	(0.022)	(0.030)	(0.024)	(0.029)
Unusually hot skin	0.078	0.084	0.088	0.089	0.084
	(0.115)	(0.121)	(0.127)	(0.120)	(0.121)
Other: rash, spots, or itch	0.006	0.011*	0.011*	0.008	0.009
	(0.037)	(0.048)	(0.041)	(0.037)	(0.041)
Other: cold symptoms	0.173	0.176	0.181	0.206**	0.184
	(0.223)	(0.220)	(0.227)	(0.235)	(0.226)
Other: ear ache	0.002	0.003	0.004	0.003	0.003
	(0.018)	(0.026)	(0.024)	(0.020)	(0.022)
Other: wound, injury, or burn	0.010	0.011	0.011	0.020***	0.013
	(0.036)	(0.044)	(0.049)	(0.061)	(0.048)
Other symptoms	0.016	0.014	0.013	0.014	0.014
	(0.047)	(0.047)	(0.051)	(0.045)	(0.048)
Mother concerned	0.066	0.079	0.057	0.057	0.065
	(0.135)	(0.136)	(0.115)	(0.102)	(0.123)
Mother very concerned	0.008	0.007	0.004**	0.007	0.007
	(0.033)	(0.028)	(0.015)	(0.020)	(0.025)
Total days observed per child	**60.115**	**59.896**	**59.403**	**60.31**	**59.929**
	(9.497)	(9.172)	(10.086)	(8.751)	(9.393)
Number of children	401	384	390	378	1553

Standard deviations in parentheses. Significance levels: *** 1%, ** 5%, * 10%, mean difference from control.

In this context, a non-urgent referral is interpreted as "within 24 hours", that is, at least on the next day. An immediate referral is interpreted to mean on the same day.

Note that cough or cold symptoms alone, while very frequent, are typically signs of a simple cold and do not require formal care. Note also that we do *not* classify children with wounds or injuries or children with pain symptoms other than ear pain to require formal care. These are symptoms that were reported frequently but are not mentioned in the C-IMCI guidelines. Lastly, we classified Sikoro as a low-malaria region, on the basis that malaria only occurs seasonally and the Malian government mandates malaria testing for all potential malaria cases. In high malaria regions, any fever is treated as likely malaria, regardless of accompanying symptoms and often without additional testing. In low-malaria regions a fever requires medical care at most a day later.

C Additional Estimation Results

Cox Hazard Estimates Using All Spells

Table 12: Cox model estimates using all spells, including left-censored.

COX HAZARD MODEL	(1)	(2)	(3)	(4)	(5)	(6)
Early: care not yet required (IMCI)			0.147***	0.154***	0.141***	0.162***
			(0.0185)	(0.0517)	(0.0191)	(0.0557)
Healthworker visits (HW)	0.972	0.854	0.895	0.868	0.779	0.749
	(0.207)	(0.190)	(0.188)	(0.199)	(0.173)	(0.179)
Early x HW (EHW)				1.224		1.303
				(0.552)		(0.617)
Free care (FC)	3.047***	2.792***	2.953***	3.064***	2.639***	2.761***
	(0.490)	(0.452)	(0.472)	(0.538)	(0.428)	(0.500)
Early x FC (EFC)				0.787		0.752
				(0.305)		(0.299)
HW and FC (HWFC)	0.867	1.057	0.952	0.939	1.209	1.238
	(0.218)	(0.279)	(0.240)	(0.257)	(0.324)	(0.357)
Early x HWFC (EHWFC)				1.083		0.858
				(0.583)		(0.488)
Total assets in US$ (log)		1.051**			1.052**	1.052**
		(0.0232)			(0.0229)	(0.0229)
Dist. to closest formal (log)		0.913			0.872	0.873
		(0.0973)			(0.0911)	(0.0916)
Child controls	-	YES	-	-	YES	YES
Household head controls	-	YES	-	-	YES	YES
Observations	24,741	21,725	24,741	24,741	21,725	21,725
Partial LL	-3617	-3165	-3455	-3454	-3017	-3016
Wald chi-square	73.09	103.7	302.7	330.6	303.4	334.6
p-value EFC = EHWFC				0.710		0.882

Significance levels: *** p<0.01, ** p<0.05, * p<0.1.
Standard errors clustered at the compound level (in parentheses).

Specification Check Proportional Hazard Model

If the hazard function is indeed proportional, the graph of $-\log(-\log(S(t,x))$ against the log of time should be parallel for different covariate values. We can check this by graphing the transformed empirical survival curves in the set of all spells for each treatment group. Since treatment assignment was random, initially the unconditional covariate distribution in the four groups should be similar. The test is imperfect if there is heterogeneity in the hazard of care for different population subgroups, since this means that the population composition will change differentially over time. Nonetheless, it can give us some idea how restrictive the proportionality assumption is in our data.

Figure C.1: Predicted survivor functions by treatment group

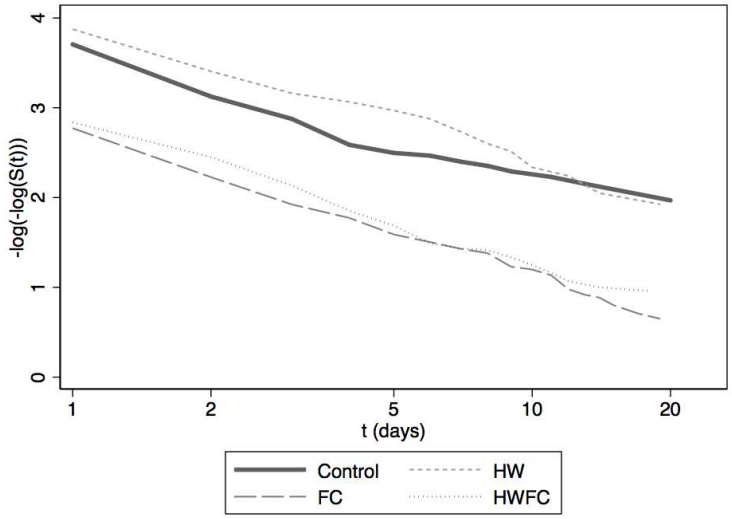


Figure C.1 shows the plot of the transformed survival function against time in log scale. The FC and HWFC curves are nearly identical. The HW and control curves are roughly shifted up in parallel, although the control group seems to show a slow-down of the survival rate at around the fourth spell day, relative to the other groups. Overall, however, the proportionality assumption seems to provide a fairly good approximation of the data, since the difference in slope across control and treatment is small in comparison to the magnitude of the downward shift in $S(t)$ in the groups that receive free care. Before day 10, the HW group also shows higher survival levels than the control group, consistent with the idea that the healthworkers delay care seeking.

Cox Hazard Estimates Using Only Uncensored Spells

Table 13: Cox estimates for uncensored spells only.

COX HAZARD MODEL	(1)	(2)	(3)	(4)	(5)	(6)
Early: care not yet required (IMCI)			0.397***	0.472***	0.398***	0.424***
			(0.0537)	(0.133)	(0.0580)	(0.120)
Healthworker visits (HW)	0.693*	0.585**	0.680**	0.688*	0.550***	0.533***
	(0.130)	(0.125)	(0.123)	(0.131)	(0.117)	(0.118)
Early x HW (EHW)				0.957		1.324
				(0.460)		(0.739)
Free care (FC)	0.862	0.802	0.891	0.955	0.795*	0.828
	(0.124)	(0.122)	(0.113)	(0.122)	(0.106)	(0.117)
Early x FC (EFC)				0.659		0.785
				(0.244)		(0.307)
HW and FC (HWFC)	1.345	1.629**	1.364	1.262	1.709**	1.692**
	(0.301)	(0.403)	(0.295)	(0.292)	(0.421)	(0.443)
Early x HWFC (EHWFC)				1.529		0.974
				(0.890)		(0.651)
Total assets in US$ (log)		1.045**			1.035*	1.035*
		(0.0232)			(0.0229)	(0.0229)
Dist. to closest formal (log)		0.948			0.927	0.932
		(0.0775)			(0.0700)	(0.0702)
Child controls	-	YES	-	-	YES	YES
Household head controls	-	YES	-	-	YES	YES
Observations	1,969	1,750	1,969	1,969	1,750	1,750
Partial LL	-2134	-1835	-2106	-2105	-1811	-1810
Wald chi-square	4.552	24.38	55.39	60.73	69.98	74.69
p-value EFC = EHWFC				0.332		0.822

Significance levels: *** p<0.01, ** p<0.05, * p<0.1.
Standard errors clustered at the compound level (in parentheses).

Cost of a Formal Healthcare Consultation

Table 14: The cost of a formal consultation by spell day, spell day X Early, and treatment group. The average "Early" effect on formal care costs is CFA 1803 (significant at 1% level in a regression on spell day, treatment, and one "Early" indicator).

	Total cost of formal visit					
	All formal		Private care only		CSCOM care only	
Healthworker visits (HW)	41.56		-1,677		0.718	
	(773.3)		(2,171)		(651.1)	
Free care (FC)	-581.3		-441.2		-164.6	
	(628.5)		(2,631)		(413.4)	
HW and FC (HWFC)	-381.3		2,653		-217.9	
	(864.0)		(3,490)		(741.8)	
Spell day	Main	x Early Care:	Main	x Early Care:	Main	x Early Care:
Day 1	4,948***	-1,348***	8,424***	-4,118	4,208***	-1,215***
	(696.9)	(507.8)	(2,641)	(2,502)	(418.0)	(452.0)
Day 2	5,445***	-1,169	9,216***	-4,886	4,762***	-915.6
	(763.7)	(760.8)	(3,170)	(5,220)	(548.3)	(657.2)
Day 3	5,499***	-2,156***	10,600***	-5,865*	4,537***	-1,802***
	(753.0)	(704.8)	(2,937)	(2,937)	(616.9)	(661.1)
Day 4	4,057***	-582.4	4,853	-	3,542***	-532.6
	(670.7)	(646.8)	(3,499)	-	(481.8)	(551.7)
Day 5	6,681***	-	14,706***	-	5,407***	-
	(945.6)	-	(2,971)	-	(757.1)	-
Day 6	6,381***	-3,655***	-	-	5,908***	-3,348***
	(1,043)	(1,273)	-	-	(937.6)	(1,222)
Day 7	5,564***	-4,564***	1,000	-	5,248***	-
	(1,111)	(1,111)	(0)	-	(1,015)	-
Day 8	5,138***	-2,557***	6,130***	-	4,891***	-2,726**
	(740.3)	(645.4)	(1,695)	-	(1,268)	(1,229)
Day 9	6,490***	-3,812***	-	-	6,163***	-3,943***
	(1,209)	(1,325)	-	-	(1,144)	(1,330)
Day 10	4,889***	-1,784	5,677**	-	4,748***	-1,892
	(757.3)	(2,005)	(2,171)	-	(742.3)	(1,879)
>10 days	4,795***	-887.5	3,303	-272.4	4,625***	357.2
	(761.6)	(1,334)	(2,645)	(3,247)	(638.7)	(563.5)
Observations	414		46		368	
R-squared	0.664		0.748		0.692	