DISCUSSION PAPER SERIES

Evaluating Nationwide Health Interventions When Standard Before-After Doesn't Work: Malawi's ITN Distribution Program

Eva Deuchert
Conny Wunsch

I Z A

Evaluating Nationwide Health Interventions When Standard Before-After Doesn't Work: Malawi's ITN Distribution Program

Eva Deuchert
CDI, University of St. Gallen

Conny Wunsch
SEW, University of St. Gallen,
CESifo and IZA

ABSTRACT

Evaluating Nationwide Health Interventions When Standard Before-After Doesn't Work: Malawi's ITN Distribution Program[*]

Nationwide health interventions are difficult to evaluate as contemporaneous control groups do not exist and before-after approaches are usually infeasible. We propose an alternative semi-parametric estimator that is based on the assumption that the intervention has no direct effect on the health outcome but influences the outcome only through its effect on individual behavior. We show that in this case the evaluation problem can be divided into two parts: (i) the effect of the intervention on behavior, for which a conditional before-after assumption is more plausible; and (ii) the effect of the behavior on the health outcome, where we exploit that a contemporaneous control groups exists for behavior. The proposed estimator is used to evaluate one of Malawi's main malaria prevention campaigns, a nationwide insecticide-treated-net (ITN) distribution scheme, in terms of its effect on infant mortality. We exploit that the program affects child mortality only via bed net usage. We find that Malawi's ITN distribution campaign reduced child mortality by 1 percentage point, which corresponds to about 30% of the total reduction in infant mortality over the study period.

Corresponding author:

Conny Wunsch
Swiss Institute for Empirical Economic Research
University of St. Gallen
Varnbüelstrasse 14
9000 St. Gallen
Switzerland
E-mail: conny.wunsch@unisg.ch

---

1 Introduction

Nationwide public health interventions are common and can be quite costly.[1] Despite of this, rigorous evaluations of the effectiveness of such interventions are rarely undertaken, the main reason being a methodological one. If the intervention is nationwide, no contemporaneous control group exists. The standard approach in such a case would be to use a before-after estimator. However, this method requires the absence of any other (uncontrolled) influence factors that changed at the time of the intervention (Heckman et al., 1999). This is difficult to argue in any application, however, in the case of health interventions the before-after method seems particularly infeasible. Countries usually implement several health interventions at the same time, and environmental hazards usually impose common trends in health risks.

If the intervention is implemented stepwise in different regions, an alternative approach would be difference-in-differences (DiD) that exploits this variation (e.g. Armecin et al., 2006; Wagstaff, 2007). However, in the case of health interventions a common problem is that many diseases have strong geographical patterns with the respective regions being subject to different environmental and also economic trends which invalidates the DiD assumptions.

We propose an alternative method to evaluate nationwide health interventions when neither the standard before-after estimator nor the difference-in-differences approach are feasible. We develop an estimator that is suitable if the effect of the health intervention on the health outcome only works through its effect on a health seeking behavior. Anti-smoking campaigns, for example, affect health only via their effect on smoking but do not animate people to eat healthier or to practice safer sex. We show that in this case one can split the total intervention effect into two components: the effect of the intervention on the health behavior (for example smoking, sexual behavior, or the usage of insecticide-treated bed nets), and the effect of the health behavior on the health outcome for those who changed behavior in reaction to the intervention.

We non-parametrically identify these two effects based on different strategies. Firstly, to identify the effect of the intervention on behavior we impose the before-after assumption on behavior rather than on the health outcome conditional on all relevant time varying factors that affect behavior. We argue that health seeking behavior is much more stable and independent of, for example, environmental trends than health outcomes. Moreover, potentially time changing factors that impact

---

[1] In the U.S., for example, tobacco control expenditures were 1.3 billion USD in 2001/2002 (Marlow, 2008).

on health behavior are much more easily observed and can thus be controlled for. Secondly, having available very informative data in our application, we rely on the conditional independence assumption (or selection on observables; Rubin, 1974) to identify the effect of behavior on the health outcome by conditioning on all factors that jointly determine behavior and the health outcome (though other approaches could in principle be used as well). The key advantage here is that a contemporaneous control group exists for behavior as not everyone adopts the behavior. We propose a semi-parametric inverse-probability-weighting estimator to estimate the average effect of the intervention and show that it is consistent and asymptotically normally distributed.

We apply the proposed estimator to evaluate the effectiveness of a national insecticide-treated-net (ITN) distribution campaign that is part of Malawi's Roll Back Malaria Initiative. Malaria is among the leading causes of infant death in Sub-Saharan Africa (Bryce et al., 2005), and it is widely believed that it causes an enormous economic burden (Bloom and Sachs, 1998; Sachs and Malaney, 2002). Substantial funds are spent on malaria control efforts, the focus lying on prevention, in particular on the provision of ITNs (WHO, 2008). Malawi's campaign is a very interesting one to study because it was the first national ITN distribution program in Sub-Saharan Africa, thus serving as a role model for many other countries.

There are few attempts to evaluate the effectiveness of net distribution schemes in terms of net coverage and usage (Cohen and Dupas, 2008; Dupas, 2009) but the effectiveness of those policies in terms of health outcomes (malaria prevalence, infant death, etc.) is unclear. We evaluate Malawi's ITN distribution campaign in terms of averted infant deaths as children are the most vulnerable to malaria. We argue that our proposed identification strategy is appropriate in this case, as the net distribution campaign affects child mortality only via its effect on bed net usage. We find that Malawi's ITN distribution campaign reduced infant mortality by 1 percentage point, which corresponds to about 30% of the total reduction in infant mortality over the study period.

The remainder of the paper is organized as follows. Section 2 develops the econometric framework. Section 3 contains the application to Malawi's ITN distribution program as part of its Roll Back Malaria Initiative. The final section concludes.

2 Econometric framework

2.1 Identification

We are interested in the effect of a nationwide intervention I (Malawi's ITN distribution program) on an outcome variable Y (child mortality). Using the notation of the potential outcome framework of the evaluation literature (Rubin, 1974, 1978), we denote by Y_t^1 the potential outcome if subject to the intervention ($I = 1$), and by Y_t^0 the potential outcome without the intervention ($I = 0$). The subscript t refers to the time period when the outcome is measured. The observed counterparts are denoted by $Y_t \equiv (1 - I)Y_t^0 + IY_t^1$. Our parameter of interest is the average treatment effect on the treated measured after the intervention at $t = 1$. Since everyone is treated in our application, this is equivalent to the average treatment effect (ATE) and the intention-to-treat effect:

$$E(Y_1^1 - Y_1^0|I = 1) = E(Y_1^1 - Y_1^0) \equiv ATE_1. \tag{1}$$

A natural choice for evaluating such universal interventions where contemporaneous comparison groups are not available because everybody is subject to the intervention would be the before-after estimator. In its conventional form, the before-after estimator assumes that the pre-intervention health outcome (at $t = 0$) proxies the potential no-intervention health outcome in the post-intervention period (at $t = 1$), i.e. $E(Y_0^0) = E(Y_1^0)$ (Heckman et al., 1999). Longitudinal data is not necessary as long as individuals from the post-intervention survey are compared to similar persons from the pre-intervention survey. However, the assumption that the expected outcome in the no-intervention state is the same in the post- and the pre-intervention period is often violated because of common unobservable trends. This is particularly relevant for health outcomes, as health is usually affected by various interventions like availability of new drugs as well as environmental factors which are particularly important when diseases have strong geographical patterns.

We propose an alternative strategy to estimate the effect of a nationwide health intervention. Health is often affected by certain behaviors B such as smoking, drinking, unsafe sex, or the failure to use bed nets. Health interventions are often designed to change this behavior B. Examples are anti-smoking or bed net distribution campaigns. If the effect of the intervention on some health outcome Y only works through a change in behavior B the causal chain is $I \rightarrow B \rightarrow Y$. We show that the evaluation problem can then be decomposed into two problems: the effect of the

3

intervention on behavior, $I \to B$, and the effect of a change in behavior on the outcome, $B \to Y$.

Denote by B_t^1 and B_t^0 the potential outcomes in terms of behavior with and without the intervention, respectively. The observed counterparts are denoted by $B_t \equiv (1 - I)B_t^0 + IB_t^1$. The potential health outcomes can then be rewritten as a function of behavior-specific potential outcomes: $Y_t^I = B_t^I Y_t^{I1} + (1 - B_t^I)Y_t^{I0}$, $I \in \{0, 1\}$. We will show that under a set of assumptions the effect of interest measured at $t = 1$ can be written as

$$E(Y_1^1 - Y_1^0) = E(Y_1^{11} - Y_1^{10}|B_1^0 = 0, B_1^1 = 1)P(B_1^0 = 0, B_1^1 = 1) \tag{2}$$

where $P(B_1^0 = 0, B_1^1 = 1)$ is the fraction of individuals that change behavior due to the intervention, $I \to B$, and where $E(Y_1^{11} - Y_1^{10}|B_1^0 = 0, B_1^1 = 1)$ is the effect of a change in behavior on the outcome for those who actually change behavior as a result of the intervention, $B \to Y|I \to B$.

Using the behavior-specific potential outcomes the assumption that the effect of I on Y only works through B can be formalized as

$$(A1): \ E(Y_1^{0B}) = E(Y_1^{1B}), \ B \in \{0, 1\}. \tag{3}$$

The assumption implies that there is no effect of the intervention on the potential outcomes of those who do not change behavior. A bed net distribution campaign for example, increases bed net usage but has no other health effects such as preventing drinking or making people more likely to do sports. The second assumption we need is a monotonicity assumption that the intervention shifts the behavior of everyone in the same desired direction (from 0 to 1 in our case) so that 'perverse' effects are ruled out:

$$(A2): \ P(B_1^0 = 1, B_1^1 = 0) = 0 \Leftrightarrow B_1^1 \geq B_1^0. \tag{4}$$

Thus, people either do not change behavior as a result of the intervention, or they change behavior in the desired way. There are no people that abstain from using bed nets in reaction to the intervention when they would use bed nets without the intervention (no defiers).

In Appendix A we show that under assumptions $A1$ and $A2$ the parameter of interest can be rewritten as the effect of the intervention for those who change behavior in the intended direction (so-called compliers) times the probability to be a complier. If the intervention only works through a change in behavior, the effect for those who do not change behavior is zero. If, in addition, there

are no defiers who change behavior in the unintended direction because of the intervention, then the total effect can only come from the compliers. So we compare the outcomes for those who actually change behavior due to the intervention and weight the results by the probability to do so.[2] We also show that under assumptions $A1$ and $A2$ the effect of the intervention for those who change behavior is equal to the effect of behavior for this population, which reflects the assumption that the intervention works through a change in behavior:

$$
\begin{aligned}
ATE_1 &= E(Y_1^1 - Y_1^0 | B_1^0 = 0, B_1^1 = 1)P(B_1^0 = 0, B_1^1 = 1) \\
&= E(Y_1^{11} - Y_1^{10} | B_1^0 = 0, B_1^1 = 1)P(B_1^0 = 0, B_1^1 = 1).
\end{aligned}
\tag{5}
$$

The problem is, that the population of compliers is unobservable because we only observe $B_0 = B_0^0$ and $B_1 = B_1^1$ but not B_0^1 and B_1^0. However, if we can plausibly assume that conditional on a set of observed covariates X_1^B measured at $t = 1$, expected hypothetical no intervention behavior at $t = 1$ is the same as expected no intervention behavior at $t = 0$, we can replace the unobserved B_0^1 by the observed B_0^0:

$$
(A3): \quad P(B_1^0 = 1 | X_1^B = x^B) = P(B_0^0 = 1 | X_1^B = x^B).
\tag{6}
$$

Assumption A3 means essentially that we observe all behavior-affecting factors that differ in the populations at $t = 0$ and $t = 1$. This is essentially a conditional before-after assumption imposed on behavior rather than on the health outcome of interest. We argue that this assumption is plausible for behavior, in particular for bed net usage, but not for health outcomes since the behavior is likely to me more stable and much less dependent on or independent of time-varying factors like availability of drugs, quality of health care and environmental factors. Moreover, potential confounders are usually more easily observed. As shown in Appendix A, using the fact that there are no defiers (A2) as well as assumption A3 we can identify the conditional probability to be a complier as the difference in expected behavior between $t = 1$ and $t = 0$ conditional on a set of covariates:

$$
\begin{aligned}
P(B_1^0 = 0, B_1^1 = 1 | X_1^B = x^B) &= E(B_1^1 - B_0^0 | X_1^B = x^B) \\
&= E(B_1 | X_1^B = x^B, I = 1) - E(B_0 | X_1^B = x^B, I = 0).
\end{aligned}
\tag{7}
$$

---

[2] Thus, the proposed framework is closely related to conditional nonparametric IV estimation of local average treatment effects (LATE) with binary instrument and treatment (Frölich, 2007; Imbens and Angrist, 1994): $LATE_1 = ATE_1 / P(B_1^0 = 0, B_1^1 = 1)$. B would correspond to the treatment and I to the instrument. The health intervention works like an instrument for health behavior with respect to the health outcome, as it directly affects only the former but not the latter. However, we are interested in the effect of I rather than B.

In order to identify the remaining component $E(Y_1^{11} - Y_1^{10}|B_1^0 = 0, B_1^1 = 1)$, we need two further assumptions. If we can observe all factors X_1^Y that jointly determine behavior and the health outcome of interest at $t = 1$, potential outcomes are independent of behavior conditional on these factors (Rubin, 1974):

$$(A4): \ Y_1^{11}, Y_1^{10} \perp B_1 | X_1^Y. \tag{8}$$

Given that there is common support with respect to the distribution of X_1^Y in the $B_1 = 0$ and the $B_1 = 1$ populations:

$$(A5): \ 0 < p_1(x_1^Y) \equiv P(B_1 = 1|X_1^Y = x^Y) < 1. \tag{9}$$

we can then identify the effect of behavior on the outcome from individuals with $B_1 = 1$ and comparable individuals with $B_1 = 0$. In Appendix A we show that assumptions A1-A5 imply that the effect of interest is nonparametrically identified from observed objects:

$$
\begin{aligned}
E(Y_1^1 - Y_1^0) & \tag{10} \\
&= \int E(Y_1^{11} - Y_1^{10}|X_1 = x)E(B_1^1 - B_1^0|X_1 = x)f_{X_1}(x)dx \\
&= \int [E(Y_1|B_1 = 1, X_1 = x, I = 1) - E(Y_1|B_1 = 0, X_1 = x, I = 1)] \\
&\quad \times [E(B_1|X_1 = x, I = 1) - E(B_0|X_1 = x, I = 0)]f_{X_1}(x|I = 1)dx,
\end{aligned}
$$

where $X_1 \equiv X_1^B \cap X_1^Y$. That means that we have to condition on all variables X_1^Y that jointly affect behavior and the outcome of interest as well as those variables X_1^B that affect behavior and differ between the populations observed at $t = 0$ and $t = 1$.

As a final remark it is important to note that by using the potential outcome framework to show identification and derive our estimator we also assume that the potential outcomes are independent of actual treatment assignment, i.e. we make the stable unit treatment value assumption (SUTVA; Rubin, 1978). For health interventions the main concern are spill over effects which may change the health outcomes of nontreated because they may benefit from fewer overall infections. We will discuss this potential problem in the context of our particular application in Section 3.3.4.

2.2 Estimation

In order to derive an estimator for ATE_1, we use Bayes' rule to rewrite

$$f_{X_1}(x|I=1) = \frac{P(B_1 = b|I=1)f_{X_1}(x|B_1 = b, I=1)}{P(B_1 = b|X_1 = x, I=1)}. \tag{11}$$

Let N_1 denote the number of post-intervention observations at $t = 1$ that satisfy common support (9). Furthermore, let x_{1i}, b_{1i} and y_{1i} denote the observed values of, respectively, X_1, B_1 and Y_1 for individual i. Moreover, let $\widehat{p}_1(x_1)$ denote the estimator for $P(B_1 = 1|X_1 = x, I = 1)$, and let $\widehat{p}_0(x_0)$ denote the estimator for $P(B_0 = 1|X_1 = x, I = 0)$. To obtain $\widehat{p}_0(x_1) \equiv P(B_0^1 = 1|X_1 = x, I = 1)$ we will estimate parametrically $P(B_0 = 1|X_0 = x, I = 0)$ using the data at $t = 0$ and then calculate the predicted values for the population at $t = 1$ using the estimated coefficients. A semi-parametric estimator for ATE_1 is then given by

$$\widehat{ATE_1} = \frac{1}{N_1}\sum_{i=1}^{N_1}\left\{\left[\frac{b_{1i}}{\widehat{p}_1(x_{1i})} - \frac{1-b_{1i}}{1-\widehat{p}_1(x_{1i})}\right]\left[\widehat{p}_1(x_{1i}) - \widehat{p}_0(x_{1i})\right]\right\}y_{1i} \tag{12}$$

$\widehat{ATE_1}$ is called an inverse-probability-weighting (IPW) estimator and the weights are given by the term in brackets. Newey (1984) shows that under the standard regularity conditions used for GMM estimation, which are usually considered to be weak, an estimator of this type is $\sqrt{N}$-consistent and asymptotically normally distributed if $\widehat{p}_0(x_1)$ and $\widehat{p}_1(x_1)$ are consistent and asymptotically normal estimators of $p_0(x_1)$ and $p_1(x_1)$, respectively. Using the result by Newey (1984) we derive the asymptotic distribution of our estimator in Appendix B. Here we take into account that the coefficients needed to construct $\widehat{p}_0(x_1)$ are estimated based on pre-treatment data, while all other components are estimated based on post-treatment data.

3 Evaluating Malawi's ITN distribution scheme

3.1 The Roll-Back-Malaria Initiative

Malaria is a vector-born infection caused by protozoan parasites. It is widespread in tropical and sub-tropical regions; estimated 3.3 billion people worldwide are at risk. In 2006, Malaria has caused estimated 247 million cases and nearly one million deaths, mostly among children under the age of five. From the 109 countries with a malaria epidemic, 45 are in Sub-Saharan Africa (WHO, 2008).

People are usually infected with malaria by a bite from the female Anopheles mosquito.[3] There is currently no vaccine against malaria, but effective prophylactic drug treatments (even though not 100% effective) are available. Other possibilities to prevent malaria are the usage of (insecticide-treated) nets and indoor residual spraying. Most adults in endemic areas are (partially) immune against an infection. Pregnancy however, reduces a woman's immunity to malaria, making her more susceptible for an infection. Children belong to the most vulnerable group, because they have not yet developed resistance. Malaria symptoms include fever and flu-like illnesses. Severe malaria, if untreated, can cause coma and death. Antimalarial drugs are available, which usually result in complete recovery. The most common subscribed drug is Chloroquine but the parasites' resistance to this drug has spread widely, making this drug ineffective in many affected regions (Plowe, 2009).

To provide a globally coordinated approach to fighting Malaria, the Roll Back Malaria (RBM) Partnership was launched in 1998. With the Abuja declaration from 2000, African leaders committed themselves to halve the number of deaths caused by malaria until 2010. The declaration's goals are to guarantee access to treatment within 24h after the onset of symptoms, to protective measures such as insecticide-treated nets (ITNs), and to provide chemoprophylaxis to at least 60% of the population at risk (RBM, 2003). The effectiveness of the RBM initiative is unclear and the partnership has been criticized sharply for its loose association structure, its inadequate and conflicting advise given by the partner institutions, and, against the background of increasing resistance to drugs, its focus on monotherapies (Yamey, 2004).

Malaria is one of the major public health problems in Malawi, accounting for about 40% of all hospital deaths among children under the age of five, mainly because access to treatment is low (NSO, 2005). To control malaria, the government has implemented several strategies through the National Malaria Control Programme. The programs include a better malaria case management, and the provision of prophylactic drug treatment to pregnant women. Its main focus however, is on malaria prevention through the distribution of ITNs.

In 1998, the Government of Malawi launched a pilot social marketing program to distribute ITNs in the Blantyre district. In 2002, the program expanded to become the first national ITN program in Sub-Saharan Africa. The distribution program scheme is three-fold. Pregnant women and children under five can obtain highly subsidized ITNs (MK 50/\$0.33 per net) through clinics at which health staff retains MK 10 as an incentive to sell nets. Rural families can buy nets at a cost

---

[3] Transmission via contaminated blood products is also possible.

of MK 100/$0.66 from village health committees, NGOs, and community-level health personnel. The general population (mostly in urban areas) can obtain unsubsidized nets from the commercial sector at a cost of MK 550-780/$3.86-$5.20 per net (PMI, 2007).

Over the period of 2000 to 2004, Malawi managed to increase bed net coverage from 13.1% to 41.9% (NSO, 2001, 2005). The program has been criticized, however, for its failure to reach the poor, the group which is thought to be the most vulnerable to malaria (Mathaga and Bowie, 2007). The government of Malawi has therefore revisited its ITN distribution policy in 2006, which now includes the free distribution to pregnant women and children under the age of five through the Expanded Programme on Immunization (EPI) and antenatal health care clinics, the free distribution to the "poorest of the poor", as emergency response (particularly for HIV positives), as well as a subsidized distribution through community venues, and an unsubsidized distribution over the commercial sector.[4]

Despite improvements in the coverage of bed nets, the effectiveness of the distribution program in terms of averted malaria cases has been unclear because no decline in malaria cases and deaths reported to the WHO was observed (WHO, 2008). Various reasons could potentially explain a failure of the campaign, for example that people are not well enough informed to use these bed nets in the right way, that there is a considerable misuse of bed nets (for example to use them for fishing nets), or that only people at no or little risk benefit from the distribution scheme while people at highest risk have no access to bed nets under the scheme.

The public malaria surveillance system, however, is likely to be imprecise, as it was argued that less than 10% of the worldwide Malaria cases and deaths are being reported (Breman and Holloway, 2007). It is therefore unclear, in how far the public Malaria surveillance system can accurately portray the epidemic situation. Data from infant mortality, the group which is most vulnerable to an infection, show a strong downward trend from 10.3% in 2000 to 7.6% in 2004 (NSO, 2001, 2005). It is unclear however, to which extent the ITN distribution program has contributed to this decline, since the Government of Malawi has launched several other programs to reduce child mortality, such as vitamin A supplementation and vaccination campaigns. Since Malawi's malaria program is a role model for other countries, a detailed analysis of the effectiveness of the distribution program in terms of averted deaths is important.

---

[4] Our data covers only the period before this policy change in 2006, though.

3.2 Data

We use data from Malawi's Demographic and Health Surveys conducted in the years 2000 and 2004. These are nationally representative household surveys that provide data for a wide range of indicators in the areas of population, health, and nutrition. Both surveys systematically cluster from a list of enumeration areas defined in the 1998 Malawian Census of Population and housing (560 clusters in 2000; 522 clusters in 2004). Clusters were not identical in both surveys. Geographic location of each cluster is provided (GPS codes). From the list of eligible clusters, a sample of households is drawn (for a total of 14,213 households in 2000 and 15,091 households in 2004). All women aged 15-49 in the selected households were eligible for the individual interview. We exclude households from the Blantyre district as it was subject to the intervention already before 2000.

The individual questionnaires cover various areas, including background characteristics (age, education, religion, etc.), household characteristics, and reproductive history. The latter is used to construct recent infant mortality rates, which include the survival status of all children who had been born in the year preceding the interview (2770 children in 2000; 2523 children in 2004). This definition differs from the definition of infant mortality in the Demographic and Health Surveys, which use all births within the past 4 years to predict infant mortality. There are three reasons why we use only recent births: (1) there may be a considerable recall bias in earlier births[5], (2) information on individual and household characteristics is retrospectively not available and represent only the current status, and (3) earlier birth had not been (fully) exposed to the national bed net distribution scheme that started in 2002. Infant mortality in the year before the interview is 8.8% in 2000 and 5.2% in 2004.

This paper uses bed net ownership as indicator variable for behavior, B, through which the ITN distribution program, I, affects child mortality, Y. All households are asked whether they own a bed net. In 2000, 13% of children in the sample live in households with a bed net compared to 47% in 2004. In both surveys, mortality is higher among children who were born into households without a bed net compared with children who were born into households with a bed net (8.3% vs. 7.3% in 2000; 6.4% vs. 4.4% in 2004).[6]

The data do not provide information on other health interventions (such as vaccinations for example) for deceased children. Environmental factors, such as rainfall, are also not available.

---

[5] All data, including birth weight etc., is collected retrospectively.
[6] The remaining descriptive statistics are provided in Appendix C.

Information on the quality of health care is also missing. A conventional before-after estimation is therefore not feasible. Using the stepwise implementation of the programme to construct a difference-in-difference estimator is also not possible because malaria has a strong geographical pattern, and mortality rates in different regions are subject to very different time trends for various reason, weather and climate being one of them. In the next section we discuss why we think that the framework proposed in Section 2 is suitable for evaluating the effectiveness of the ITN distribution scheme.

3.3 Plausibility of the identification strategy

The identification strategy outlined in Section 2 requires on the one hand controlling for all variables that jointly affect child mortality and bed net ownership to make the conditional independence assumption, that is necessary to estimate the effect of bed net ownership, plausible. On the other hand, we have to control for all time-variant predictors of bed net ownership to estimate the size of the complier population. Therefore, to make the identification strategy credible, we need to understand which factors determine bed net ownership and child mortality.

3.3.1 Bed net ownership

Economic theory views health as an investment good that can be produced (Grossman, 1972). Health is produced using (curative) medical services as well as individual prevention effort as inputs. Based on the Grossman model one can derive the demand for prevention efforts. Here, the model predicts that the demand for prevention is sensitive to the price of the input and to the wage or wealth level. Interestingly, since the level of education is assumed to increase the efficiency of the health investment, less investment is necessary to maintain a given stock of health capital so that people with higher education would chose to invest less. The latter result, however, is based on the assumption that everybody has perfect information on the expected benefits of the investment. If people with a high education have better information, the reverse could be true as well. The original Grossman model does not consider uncertainty of the investment and therefore, does not consider that people face different risk profiles. If people, however, have different risks of catching a disease (for example because the disease is contagious and heavily locally concentrated), the respective risk of a bad health status could be a very relevant demand factor that explains why some people are more likely to engage in preventive actions (Geoffard and Philipson, 1996).

Relatively little empirical evidence is available on the demand for bed nets. Early willingness to pay studies find that people usually reported a positive willing to pay for bed nets which varies strongly with socioeconomic status (Onwujekwe et al., 2001; Legesse et al., 2007). Dupas (2009) however, shows in an experimental setting that the verbal commitment of investing in nets does not affect their actual investment behavior. It is therefore unclear whether or not the results from hypothetic willingness-to-pay studies can portray the investment decision in bed nets.

To understand the price sensitivity of the actual demand for bed nets, two experimental studies have been carried out (Cohen and Dupas, 2008; Dupas, 2009). Both studies find that the demand for nets is sensitive to the price, but the price elasticity of demand is very small around a price of zero. Dupas (2009) furthermore tests, whether different marketing framing (health framing that emphasizes the risk of morbidity and mortality, and a financial framing that emphasizes the financial gains that could be realized if malaria could be prevented) affect the demand for bed nets. None of the interventions had a significant effect for take-up.

Of special concern is whether or not bed nets are used in the appropriate way. It is often argued that freely distributed bed nets will be less valued and may be used for alternative purposes. The empirical evidence is mixed. Cohen and Dupas (2008) find no evidence that women who received free ITNs are less likely to use them than those who paid for their bed nets. In a study area adjacent to Lake Victoria, where nets are often provided at low costs or free of charge, Minakawa et al. (2008) in contrast observe considerable misuse of bed nets (e.g. for drying fish and fishing).

Concerning other relevant demand factors the evidence is mixed. The usage of bed nets, lack of access, the perception that nets prevent malaria, and low education are the most important predictors for bed net ownership or usage (Belay and Deressa, 2008; Pettifor et al., 2008). Dupas (2009) in contrast, finds only household wealth being associated with net purchase and usage, while other indicators, such as education or health knowledge, are not associated with net usage.

In Table 1 we display bed net ownership for different household characteristics for the pre and post-intervention periods 2000 and 2004. Bed net ownership is strongly associated with household wealth (in both periods). There is a strong increase in bed net ownership over time, with the richest and middle quintiles of the wealth distribution exhibiting the strongest increase. The strong targeting of bed net distribution becomes evident from the sharp increase in bed net ownership in rural areas. The rural population can get highly subsidized bed nets from their local health care

Table 1: Bed net ownership

	2000	2004	Difference
Wealth index			
Poorest quintile	11.6%	38.1%	26.5%
Poorer quintile	8.2%	38.8%	30.6%
Middle quintile	16.3%	53.0%	36.7%
Richer quintile	17.8%	46.1%	28.2%
Richest quintile	41.2%	79.0%	37.7%
Residence			
Rural	14.4%	70.8%	56.4%
Urban	36.8%	47.3%	10.5%
Highest education (mother)			
No education	8.8%	41.3%	32.5%
Primary	18.0%	48.2%	30.3%
Secondary, higher	51.2%	73.2%	22.0%
Distance to lake			
<50 km	25.6%	58.0%	32.4%
50-99 km	9.9%	37.7%	27.8%
100+ km	7.6%	48.4%	40.8%

providers. Mother's education is strongly associated with bed net coverage, where highest coverage is observed in families where mothers have secondary and higher education. The difference vanishes somewhat in the post-intervention period. Malaria risk can be captured by the distance to the next lake since we know that the malaria burden is highest in areas close to standing water (as mosquitoes breed in standing water). We find that people life close to a lake have highest bed net coverage, but this difference vanishes also somewhat in the post-intervention period.

3.3.2 Child mortality

About 10.5 million children under the age of five are estimated to die worldwide every year. Of the 20 countries with the highest child mortality, 19 are in Sub-Saharan Africa. Infectious and parasitic diseases are the main causes of death and represent about 60% of child deaths (WHO, 2003). Malaria is one of the biggest killers, accounting for approximately 20% of child deaths in Sub-Saharan Africa (Rowe et al., 2006; Adazu et al., 2005). Other infectious diseases such as HIV/AIDS or lower respiratory infections (Garrib et al., 2006) and severe anemia (Adazu et al., 2005) are also important causes of deaths.

Economic theory suggests that the same factors that affect health seeking behavior are also likely to affect health (Grossman, 1972). Therefore, we expect that household income, wealth, and risks are negatively associated with child mortality. Because education is assumed to increase the efficiency of the health investment, and because well educated people are more likely to engage in health seeking behavior, we expect a negative association between parental education and child mortality.

A substantial body of empirical evidence exists on the determinants of child health and mortality in developing countries. This literature identifies low socio-economic status as one of the most important predictors of poor health and mortality (Omariba et al., 2007; Minujin and Delamonica, 2003). Using comparable data from 24 countries, Minujin and Delamonica (2003) for example, show that children from families belonging to the bottom quintile of the wealth distribution are three times more likely to die before age five than children belonging to the top quintile. A higher susceptibility to communicable diseases, and a poor access to health care infrastructure (which increases the costs of assessing health care) could explain the strong disparity in child health (Galiani et al., 2005; Shi, 2000; Koenig et al., 2001; Maitra and Ray, 2004).

A further important factor of child mortality is the mother's education (Iram and Butt, 2008; Omariba et al., 2007). Mother's education can influence child survival by different ways. Education for example, can lead to a higher efficiency in producing child health, may shape mother's preferences, and raises income (Schultz, 1984). A particular emphasis has been placed on the role of maternal health knowledge (Kovsted et al., 2002; Glewwe, 1999). A study conducted in Guinea-Bissau for example, finds a strong and positive effect of health knowledge (measured by knowing malaria transmission modes and preventive actions) on child health and survival that 'crowed out' the effects of the mother's education (Kovsted et al., 2002).

Maternal education however, may not be the only factor relevant for health-related decision making. Health related skills, values, and information are public goods within a household. Therefore, considerable intra-household externalities of the education of other household members may exist. Lindelow (2008) for example shows that after controlling for maternal education and income, education of other (non-spousal) household members has a significant and large effect on health care choices such as maternity services and child immunization in Mozambique.

There are also important regional differences in the determinants of child mortality. Malaria for example, is strongly associated with the distance to the next water body (Noor et al., 2008) and with lower altitude (Bodker et al., 2003). Furthermore, a wide gap between rural and urban areas had been observed, with rural areas having a higher level and a lower reduction in child mortality compared to the urban population (Wang, 2002). These differences may be explained by low access to health care or by poverty being mainly concentrated in rural areas. Additionally, child characteristics such as a low birth weight (Guilkey and Riphahn, 1998), twin birth (van der Mei

et al., 2003), and the gender of the child (Wamani et al., 2007) are strongly associated with poor infant health and mortality.

3.3.3 Selection of control variables

To simplify the estimation we estimate probit models for $p_0(x_0) = P(B_0 = 1 | X_0 = x, I = 0)$ and $p_1(x_1) = P(B_1 = 1 | X_1 = x, I = 1)$ using one set of explanatory variables that includes both the factors that jointly determine bed net ownership and child mortality, and the predictors of bed net ownership that may differ between the pre and post-intervention survey samples. The major factors that should be included are variables that proxy the socio-economic status, education and health knowledge, access to health care and bed nets, and malaria risks. The Demographic and Health Surveys provide a large set of covariates that can be used to approximate these factors. We split these variables into five different categories: child characteristics, mother characteristics, household characteristics, partner characteristics, and regional characteristics.

Child characteristics: The characteristics of the child are clearly among the most important factors we have to control for. We observe (hypothetical) age, twin birth, sex, size at birth and place of delivery. While these factors are clearly likely to affect infant mortality they are also likely to affect health-related behavior such as bed net usage. On the one hand, adverse conditions at birth might have raised the awareness for potential health risks as well as how much the parents worry and care for the child. On the other hand, contact with health providers at birth may have increased the knowledge about health risks and their prevention. Moreover, Malawi's distribution system imposes a price discrimination, where children can get highly subsidized bed nets from their health care provider. We therefore expect that children who were born in a health facility are more likely to sleep below nets, because they pay the lowest price for these nets. Finally, the place of birth is also an important determinant of infant mortality, as the severity of birth complications can vary with the place of delivery.

Mother's characteristics: Important factors for health and health seeking behavior are the education of the mother, her knowledge and awareness of health risks and their prevention, as well as the attitude towards health-related issues. We are able to capture this by the mother's education, whether older children had died in the past which may raise a mother's awareness of health issues, health knowledge measured by knowledge on dehydration, and whether a family planning worker visited in the family in the last 12 months. Barriers that prevent women to access health care

may be also associated with child mortality and bed net usage. We therefore include self-reported barriers to access health care like needing permission, transport or money, as well as not wanting to go alone and worries about the gender of the physician. Further control variables are the age, marital status, labor market status, ethnicity, religion, and long term health (measured by mother's height) that are included to measure other aspects of knowledge, preferences and behavior.

Partner's characteristics: In addition to the mother's characteristics, her partner's characteristics may be associated with bed net ownership and mortality for similar reasons. The DHS is less informative regarding fathers. Here, we include his age and education.

Household characteristics: One of the most important factors that are likely to affect both bed net ownership and infant mortality are household wealth and income. The data include the DHS wealth index, which is a linear combination of household item indicator variables. Indicator weights are based on a principal component analysis (Rutstein and Johnson, 2004). A complication arises when using this index here. Neither the household item indicators nor the indicator weights are identical in the two waves. Therefore, the DHS wealth indices are not comparable. We re-estimate the wealth index using the following household items: electricity, radio, television, bicycle, motorcycle, car, type of toilet, type of water supply, type of house floor, and type of cooking material. The correlation coefficient between the predicted wealth index and the original index is 0.84 for wave 1 and 0.96 for wave 2. Unfortunately, the data do not provide direct information on income. However, information on employment status and education (mother and partner) as well as the household wealth index are likely to capture most of the relevant variation.

Regional characteristics: An important determinant of health and bed net ownership are variables that capture the risk to catch malaria. Besides being more exposed to the disease, the higher the risk the more likely it is that parents are aware of the risk and protective measures and thus more inclined to use them. Malaria is generally found at altitudes lower than 1600 m because temperature above this altitude restricts mosquito and parasite growth. We therefore include altitude as one proxy for malaria risk. Furthermore, water bodies are mosquito breeding sites suggesting that the distance from a major water body should be strongly associated with malaria risk. We therefore include the distance to one of the major lakes (Lake Malawi, Lake Malombe, Lake Chilwa) which estimated from GPS information.[7] Other regional characteristics such as region and the distance

---

[7] GPS boundaries for the major lakes are provided by www.sahims.net.

to the next district capital[8] are included to measure barriers to health care and as proxies for the potential availability of and access to bed nets.

3.3.4 Potential spill over effects

Another important assumption we make to identify the effect of interest is SUTVA. In our example, SUTVA requires that the observed mortality of a particular respondent depends only on his own behavior but not on the bed net usage of others around that person. In terms of an infectious disease, this assumption is often likely to be violated (Miguel and Kremer, 2004). The main transmission mechanism for Malaria are mosquito bites. When a large number of nets is distributed in one residential area, their chemical additives may help reduce the number of mosquitos in the environment. With fewer mosquitos in the environment, the chances of malaria infection may therefore be reduced for all. Hawley et al. (2003) study the effect of ITNs on nearby households in the context of a large-scale, group-randomized, controlled mortality trial in western Kenya. Their results indicate a strong protective effect of ITNs on control villages (those who lack ITNs) located within 300 meters of a treated village with ITNs. However, this effect vanishes quickly with no significant protective effect being observed for any outcome in villages located 600-899 meters from the nearest ITN village.

We acknowledge that this effect exists and potentially biases our results. However, we believe that this effect is unlikely to be very strong: First, our data comes from a cluster survey, where clusters are drawn from the census sample frame. In each cluster, all households were listed and from these lists, a systematic sample of households was drawn, with the number of households selected per cluster being inversely proportional to the size of the cluster. Thus, even though it is likely that some neighboring households were selected, it is very unlikely that the survey consist only of neighboring households. Second, spill over effects are most likely if ITNs are highly concentrated in one residential cluster or community. In the Kenyan trial studied by Hawley et al. (2003) for example, the externality was only observed if ITN coverage in treated villages exceeded 50%. In our sample we do observe some clusters with ITN coverage exceeding 50% but in more than 60% of our observed clusters in the year 2004, coverage was lower than 50%.

Another argument in favor of our analysis is that the spill over effects described above are likely to be positive: Nontreated individuals are likely to benefit from bed net usage of their neighbors. If this

---

[8] GPS codes for district capital are provided by Wikipedia.

is the case the health outcomes of the nontreated are positively affected and we would underestimate the effect of the program. Consequently, we will estimate at least a lower bound for the effect of Malawi's ITN distribution campaign.

3.4 Results

We estimate the average treatment effect defined by equation (12) in a two-step procedure. First, we estimate the propensity scores $\widehat{p_0}(x_0)$ using the 2000 wave and $\widehat{p_1}(x_1)$ using the 2004 wave. We use standard Probit models, with the indicator that the household owns at least one bed net as the dependent variable. Most variables are categorical variables which we include as dummy variables. For the few continuous variables (age, distances, altitude, wealth score, height) we include both the continuous variables to capture linear dependencies, and dummy variables to capture potential non-linearities. Response to the questions was very high (usually less than 1% missing, at most 5%; see Appendix C). We tested for zero association of missing answers with bed net usage, which cannot be rejected at conventional significance levels for all categories. We therefore code the missing answers as part of the reference group. All estimations adjust for survey weights. The main results are presented in Table 2. For the complete specification and results see Appendix D. The Probit specifications have been tested extensively against mis-specification in terms of omitted variables, functional form, heteroscedasticity and normality.

Among the strongest predictors for bed net ownership are the socio-economic status (measured by the household wealth as well as education of the mother and her partner). In both waves, variables that capture malaria risk (altitude and distance to a lake) are strongly associated with bed net ownership. This risk depending behavior is in line with economic theory of health seeking behavior, which should be highest if the risk imposed is highest (Geoffard and Philipson, 1996). In the post-intervention period, ownership of bed nets is strongly associated with giving birth in a health facility and living in an urban rather than rural area. This is expected as health facilities and village health committees are the key distribution channels for subsidized bed nets. Regions in general play a strong role for bed net ownership in both waves. Barriers that prevent women to access health care are significant in the post-intervention period only, which may also be explained by the formal health care sector being an important distribution channel for bed nets after the policy got introduced. Most other control variables are insignificant and do not play a major role for bed net ownership.

Table 2: Main results from Probit estimation

	Pre-intervention (2000)			Post-intervention (2004)		
	coeff.		t-stat.	coeff.		t-stat.
Constant	-0.4687		-0.56	0.9579	*	1.66
Child characteristics						
Female	0.0421		0.49	0.0319		0.50
Age (years)	0.1160		0.78	0.0398		0.38
Size at birth (reference: average, N/A)						
Very large	0.0276		0.28	-0.0433		-0.61
Very small	0.0270		0.23	-0.1166		-1.28
Place at birth (reference: home, other, N/A)						
Public hospital	0.1050		1.10	0.3429	***	4.42
Private hospital	0.0477		0.38	0.2858	***	3.48
Mother characteristics						
Age	-0.0099		-0.39	-0.0114		-0.58
Highest education (reference: primary)						
No education	-0.4323	***	-4.13	-0.2448	***	-2.93
Secondary, higher	0.4804	***	3.05	0.3806	***	2.70
Knowledge of oral rehydration (reference: not known, N/A)						
Heard of	0.1838		1.38	0.0101		0.07
Used	0.2498	*	1.85	0.0185		0.22
Labor market status: not working	0.0115		0.13	-0.0265		-0.38
Problems to get medical help (reference: no problem, N/A)						
Where to go	-0.1544		-1.05	-0.1367		-1.45
Permission to go	-0.0968		-0.56	0.0423		0.31
Money	0.0149		0.18	-0.0831		-1.08
Taking transport	0.0386		0.46	-0.1612	**	-2.24
Not wanting to go alone	-0.0011		-0.01	0.1372	*	1.72
Concern no female provider	0.1824		1.37	-0.1736	*	-1.85
Visited by a family worker in past 12 months	-0.0301		-0.27	0.0493		0.49
Any older children died	0.1176		0.88	0.1567		1.37
Mother's height (cm)	-0.0002		-0.62	0.0001		1.37
Regional characteristics						
Altitude (m)	-0.0015	***	-4.26	-0.0017	***	-6.32
Distance to lake (km)	-0.0100	***	-2.71	-0.0067	***	-2.80
Region (reference: Southern)						
Northern	0.8275	***	4.51	0.3492	**	2.24
Central	0.3212	**	2.25	0.1866		1.60
Urban area	0.0575		0.39	-0.4508	***	-3.11
Distance to next city	0.0135		1.60	0.0010		0.17
Partner characteristics						
Has no partner	0.7657	*	1.94	-0.1382		-0.41
Highest education (reference: primary, N/A)						
No education	-0.1381		-1.04	-0.1177		-1.12
Secondary, higher	0.5236	***	4.61	0.2456	***	2.58
Partners age	0.0212	*	1.95	-0.0004		-0.04
Household characteristics						
Wealth index	0.5937	***	6.37	0.4566	**	2.55

Note: ***/**/* indicates significance on the 1/5/10% level. Additional control variables: age dummies for mother and partner, ethnicity, religion, number of children ever born, dummies for mother's height, dummies for altitude, distance to lake and distance to next city, dummies for wealth quintiles (see Appendix D).

In Table 3 we present our estimates of the average effect of Malawi's ITN distribution program on child mortality. Our baseline results presented in line (0) are based on the probit estimates discussed above where we impose common support (assumption A5) by excluding all children from households without (with) bed nets with post-intervention propensity scores below the lowest (above the highest) score of those children from households with (without) bed nets. We do not find evidence

for a significant common support problem as there are only 73 children (2.9%) with propensity scores outside the common support region. We find that Malawi's ITN distribution program reduced infant mortality by about 1 percentage point. Given the total decline in child mortality of 3.6 percentage points between 2000 and 2004, the analysis suggests that the bed net distribution program caused about 30% of the total decline in the study period. The effect is at the margin of significance on the 5% level.

Table 3: Estimation results

		ATE	Bootstrap SE	Confidence interval (95%)		Obs.
(0)	Baseline	-.0105	.0055	-.0220	.0000	2450
(1)	Without support enforcement	-.0105	.0055	-.0218	.0000	2523
(2)	Support within 10th largest nontreated and 10th smallest treated $\hat{p}_1(x_{1i})$	-.0110	.0054	-.0223	-.0007	2295
(3)	Trim weights to a maximum of 1/0.1	-.0103	.0052	-.0212	.0001	2450
(4)	Trim weights to a maximum of 1/0.05	-.0105	.0054	-.0218	.0000	2450
(5)	Trim weights to a maximum of 1/0.01	-.0105	.0055	-.0220	.0000	2450
(6)	Discard observations with weight $> 1/0.1$	-.0098	.0051	-.0201	.0003	2439
(7)	Discard observations with weight $> 1/0.05$	-.0104	.0054	-.0216	.0000	2449
(8)	Discard observations with weight $> 1/0.01$	-.0105	.0055	-.0220	.0000	2450
(9)	Without sample weights	-.0090	.0045	-.0183	-.0006	2450
(10)	Without child variables	-.0107	.0054	-.0225	-.0001	2450
(11)	Without mother variables	-.0116	.0050	-.0219	-.0024	2450
(12)	Without regional variables	-.0091	.0055	-.0205	.0009	2450
(13)	Without partner variables	-.0113	.0053	-.0221	-.0012	2450
(14)	Without household variables	-.0083	.0051	-.0186	.0018	2450

Note: Bootstrap standard errors and confidence levels are based on 1999 replications. Analytical standard errors are not used because they do not take into account adjustments in the sample or the weights due to common support enforcement, trimming and discarding observations with large weights.

We perform a number of sensitivity checks to assess the robustness of our results. The first ones presented in lines (1)-(9) in Table 3 concern the implementation of the estimator. The effects on the size of the estimation sample are displayed in Table 6 in Appendix E. We start by varying the common support (CS) criterion. Both ignoring potential support problems in line (1) and implementing a stricter support criterion leaves the results essentially unchanged. In a second step we check the sensitivity of our results to observations with large weights in the inverse probability weighting. We either trim weights to a maximum of 1/0.1, 1/0.05 or 1/0.01, or discard observations with higher weights. As only very few observations are affected by this (11, 1, 0, respectively), the results remain unchanged. The results are also largely insensitive to the failure of applying survey weights in line (9). The effect becomes somewhat smaller but is estimated more precisely.

The second set of checks presented in lines (10)-(14) in Table 3 refers to our identifying assumptions. The identification strategy requires controlling for all time-varying variables that affect bed net ownership, as well as for all confounders that jointly determine bed net ownership and infant

mortality. These assumptions are not testable. However, we can analyze the sensitivity of the results to omitting relevant variables leaving blocks of covariates out of the estimation procedure. Our results are fairly robust to omitting variables in size and significance. The strongest changes occur when household or regional variables are excluded, but the estimated intervention effect is still comparable to the one that controls for all variables. The insensitiveness to omitted variables can be either explained by the correlation of control variables and the flexibility of our specification so that other (included) variables pick up the effect, or by the small variability of child mortality with respect to these omitted characteristics. As also argued by Altonji et al. (2005), this should put faith in our results, since one can assume that other unobserved characteristics are also likely to play a negligible role for predicting mortality (in particular conditional on the other characteristics).

4 Conclusion

We propose a framework to assess the effectiveness of nationwide health interventions where no contemporaneous control group is available and where the standard before-after assumption is implausible. Our identification strategy requires that the intervention is effective only via its effect on health seeking behavior (such as bed net usage in our example). If this assumption is valid, the total effect of the intervention can be split into two components: the effect of the intervention on the health seeking behavior, and the effect of the behavior on the health outcome. We show how these two effects can be identified using different identification strategies that are likely to hold in this type of applications when good data is available. We propose a semi-parametric estimator for the average effect of the intervention that is consistent and asymptotically normally distributed.

We apply this estimator to data from Malawi to evaluate the effectiveness of Malawi's insecticide-treated-net distribution program as part of its Roll Back Malaria Initiative. The results suggest that the distribution program contributed to about 30% of the total reduction in infant mortality over the study period. Acknowledging the possibility that there may be positive spill over effects of the policy, our estimate can be interpreted as a lower bound on the overall effect.

We believe that the method proposed and applied in this paper may not only be useful for evaluating this particular program, but also for evaluating other nationwide health interventions that aim at affecting health outcomes via changes in health seeking behavior, such as safer-sex or anti-tobacco campaigns.

References

ADAZU, K., K. A. LINDBLADE, D. H. ROSEN, F. ODHIAMBO, P. OFWARE, J. KWACH, A. M. V. EIJK, K. M. DECOCK, P. AMORNKUL, D. KARANJA, J. M. VULULE and L. SLUTSKER (2005): Health and demographic surveillance in rural western Kenya: a platform for evaluating interventions to reduce morbidity and mortality from infectious diseases, *The American Journal of Tropical Medicine and Hygiene*, 73(6), pp. 1151–1158.

ALTONJI, J., T. ELDER and C. TABER (2005): Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools, *Journal of Political Economy*, 113(151-184).

ARMECIN, G., J. R. BEHRMAN, P. DUAZO, S. GHUMAN, S. GULTIANO, E. M. KING and N. LEE (2006): Early Childhood Development through an Integrated Program: Evidence from the Philippines. World Bank Policy Research Working Paper 3922, May 2006, Impact Evaluation Series No. 2.

BELAY, M. and W. DERESSA (2008): Use of insecticide treated nets by pregnant women and associated factors in a pre-dominantly rural population in northern Ethiopia, *Tropical Medicine & International Health: TM & IH*, 13(10), pp. 1303–1313.

BLOOM, D. E. and J. D. SACHS (1998): Geography, Demography, and Economic Growth in Africa, *Brookings Papers on Economic Activity*, 29(1998-2), pp. 207–296.

BODKER, R., J. AKIDA, D. SHAYO, W. KISINZA, H. MSANGENI, E. PEDERSEN and S. LINDSAY. (2003): Relationship between altitude and intensity of malaria transmission in the Usambara Mountains, Tanzania, *Journal of Medical Entomology*, 40(5), pp. 706–17.

BREMAN, J. G. and C. N. HOLLOWAY (2007): Malaria Surveillance Counts, *American Journal of Tropical Medicine and Hygiene*, 77(6), p. 36–47.

BRYCE, J., C. BOSCHI-PINTO, K. SHIBUYA and R. E. BLACK (2005): WHO estimates of the causes of death in children, *Lancet*, 365(9465), pp. 1147–1152.

COHEN, J. and P. DUPAS (2008): Free Distribution or Cost-Sharing? Evidence from a Malaria Prevention Experiment. NBER Working Paper No. 14406.

DUPAS, P. (2009): What Matters (and What Does Not) in Households' Decision to Invest in Malaria Prevention?, *American Economic Review*, 99(2), pp. 224–30.

FRÖLICH, M. (2007): Nonparametric IV estimation of local average treatment effects with covariates, *Journal of Econometrics*, 139, p. 35–75.

GALIANI, S., P. GERTLER and E. SCHARGRODSKY (2005): Water for Life: The Impact of the Privatization of Water Services on Child Mortality, *Journal of Political Economy*, 113(1), pp. 83–120.

GARRIB, A., S. JAFFAR, S. KNIGHT, D. BRADSHAW and M. BENNISH (2006): Rates and causes of child mortality in an area of high HIV prevalence in rural South Africa, *Tropical Medicine & International Health*, 11(12), pp. 1841–1848.

GEOFFARD, P.-Y. and T. PHILIPSON (1996): Rational Epidemics and Their Public Control, *International Economic Review*, 37(3), pp. 603–24.

GLEWWE, P. (1999): Why Does Mother's Schooling Raise Child Health in Developing Countries? Evidence from Morocco., *Journal of Human Resources*, 34(1), pp. 124–59.

GROSSMAN, M. (1972): On the Concept of Health Capital and the Demand for Health, *Journal of Political Economy*, 80(2), pp. 223–55.

GUILKEY, D. K. and R. T. RIPHAHN (1998): The determinants of child mortality in the Philippines: estimation of a structural model, *Journal of Development Economics*, 56(2), pp. 281–305.

HANSEN, L. (1982): Large Sample Properties of Generalized Method of Moments Estimators, *Econometrica*, 50, pp. 1029–1055.

HAWLEY, W. A., P. A. PHILLIPS-HOWARD, F. O. TER KUILE, D. J. TERLOUW, J. M. VULULE, M. OM-BOK, B. L. NAHLEN, J. E. GIMNIG, S. K. KARIUKI, M. S. KOLCZAK and A. W. HIGHTOWER (2003): Community-Wide Effects of Permethrin-Treated Bed Nets on Child Mortality and Malaria Morbidity in Western Kenya, *The American Journal of Tropical Medicine and Hygiene*, 68(4 Suppl), pp. 121–127.

HECKMAN, J., R. LALONDE and J. SMITH (1999): The Economics and Econometrics of Active Labor Market Programs, in: O. Ashenfelter and D. Card (editors), *Handbook of Labour Economics*, volume 3, pp. 1865–2097, North-Holland, Amsterdam.

IMBENS, G. W. and J. D. ANGRIST (1994): Identification and Estimation of Local Average Treatment Effects, *Econometrica*, 62(2), pp. 467–75.

IRAM, U. and M. S. BUTT (2008): Socioeconomic determinants of child mortality in Pakistan: Evidence from sequential probit model, *International Journal of Social Economics*, 35(1), pp. 63–76.

KOENIG, M. A., D. BISHAI and M. A. KHAN (2001): Health Interventions and Health Equity: The Example of Measles Vaccination in Bangladesh, *Population and Development Review*, 27(2), pp. 283–302.

KOVSTED, J., C. C. POERTNER and F. TARP (2002): Child Health and Mortality: Does Health Knowledge Matter?, *Journal of African Economies*, 11(4), pp. 542–560.

LEGESSE, Y., A. TEGEGN, T. BELACHEW and K. TUSHUNE (2007): Households willingness to pay for long-lasting insecticide treated nets in three urban communities of Assosa Zone, western Ethiopia, *Ethiopian Medical Journal*, 45(4), pp. 353–362.

LINDELOW, M. (2008): Health as a Family Matter: Do Intra-household Education Externalities Matter for Maternal and Child Health?, *The Journal of Development Studies*, 44(4), pp. 562–585.

MAITRA, P. and R. RAY (2004): The Impact of Resource Inflows on Child Health: Evidence from Kwazulu-Natal, South Africa, 1993-98, *Journal of Development Studies*, 40, pp. 78–114.

MARLOW, M. L. (2008): Determinants of state tobacco-control expenditures, *Applied Economics*, 40(7), pp. 831–839.

MATHAGA, D. and C. BOWIE (2007): Malaria control in Malawi: are the poor being served?, *International Journal for Equity in Health*, 6(22).

VAN DER MEI, J., T. HEIJSTRA and E. R. BOERSMA (2003): Growth patterns and survival rates of twins in a rural area of Ghana: a follow-up study from birth to adulthood, *Annals of Tropical Paediatrics: International Child Health*, 23, pp. 107–120.

MIGUEL, E. and M. KREMER (2004): Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities, *Econometrica*, 72(1), pp. 159–217.

MINAKAWA, N., G. DIDA, G. SONYE, K. FUTAMI and S. KANEKO (2008): Unforeseen misuses of bed nets in fishing villages along Lake Victoria, *Malaria Journal*, 7(1), p. 165.

MINUJIN, A. and E. DELAMONICA (2003): Mind the Gap Widening Child Mortality Disparities, *Journal of Human Development*, 4, pp. 397–418.

NEWEY, W. (1984): A Method of Moments Interpretation of Sequential Estimators, *Economics Letters*, 14, p. 201–206.

NEWEY, W. and D. MCFADDEN (1994): Large Sample Estimation and Hypothesis Testing, in: R. Engle and D. McFadden (editors), *Handbook of Econometrics*, volume 4, pp. 2113–2245, North-Holland, Amsterdam.

NOOR, A., A. CLEMENTS, P. GETHING, G. MOLONEY, M. BORLE, T. SHEWCHUK, S. HAY and R. SNOW (2008): Spatial prediction of Plasmodium falciparum prevalence in Somalia, *Malaria Journal*, 7(1), p. 159.

NSO (2001): Malawi Demographic Health Survey 2000. Calverton, Maryland USA: National Statistical Office and ORC Macro, 2001.

NSO (2005): Malawi Demographic Health Survey 2004. Calverton, Maryland USA: National Statistical Office and ORC Macro, 2005.

OMARIBA, D., R. BEAUJOT and F. RAJULTON (2007): Determinants of infant and child mortality in Kenya: an analysis controlling for frailty effects, *Population Research and Policy Review*, 26(3), pp. 299–321.

ONWUJEKWE, O., R. CHIMA, E. SHU, D. NWAGBO and P. OKONKWO (2001): Hypothetical and actual willingness to pay for insecticide-treated nets in five Nigerian communities, *Tropical Medicine & International Health: TM & IH*, 6(7), pp. 545–553.

PETTIFOR, A., E. TAYLOR, D. NKU, S. DUVALL, M. TABALA, S. MESHNICK and F. BEHETS (2008): Bed net ownership, use and perceptions among women seeking antenatal care in Kinshasa, Democratic Republic of the Congo (DRC): Opportunities for improved maternal and child health, *BMC Public Health*, 8(1), p. 331.

PLOWE, C. (2009): The evolution of drug-resistant malaria, *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 103(1), pp. S11–S14.

PMI (2007): Malaria Operational Plan MALAWI FY 2008. Washington USA: President's Malaria Initiative, 2007.

RBM (2003): The Abuja Declaration and the Plan of Action. WHO/CDS/RMB/2003.46, Geneva: Roll Back Malaria/World Health Organization, 2003.

ROWE, A. K., S. Y. ROWE, R. W. SNOW, E. L. KORENROMP, J. R. A. SCHELLENBERG, C. STEIN, B. L. NAHLEN, J. BRYCE, R. E. BLACK and R. W. STEKETEE (2006): The burden of malaria mortality among African children in the year 2000, *International Journal of Epidemiology*, 35(3), pp. 691–704.

RUBIN, D. (1974): Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies, *Journal of Educational Psychology*, 66(5), pp. 688–701.

RUBIN, D. (1978): Bayesian Inference for Causal Effects: The Role of Randomization, *Annals of Statistics*, 6, pp. 34–58.

RUTSTEIN, S. and K. JOHNSON (2004): The DHS Wealth Index. DHS Comparative Reports No. 6, Calverton, Maryland USA.

SACHS, J. and P. MALANEY (2002): The economic and social burden of malaria, *Nature*, 415(6872), pp. 680–685.

SCHULTZ, P. (1984): Studying the impact of household economics and community variables on child mortality, in: W. Mosley and L. C. Chen (editors), *Child survival: strategies for research*, *Population and Development Review*, volume 10, pp. 215–35, Population Council, New York.

SHI, A. (2000): How access to urban potable water and sewerage connections affects child mortality. The World Bank.

WAGSTAFF, A. (2007): Health Insurance for the Poor: Initial Impacts of Vietnam's Health Care Fund for the Poor. World Bank Policy Research Working Paper 4134, February 2007, Impact Evaluation Series No. 11.

WAMANI, H., A. N. ASTRØM, S. PETERSON, J. K. TUMWINE and T. TYLLESKÄR (2007): Boys are more stunted than girls in sub-Saharan Africa: a meta-analysis of 16 demographic and health surveys, *BMC Pediatrics*, 7, p. 17.

WANG, L. (2002): Health outcomes in poor countries and policy options : empirical findings from demographic and health surveys, Technical report, The World Bank, URL http://ideas.repec.org/p/wbk/wbrwps/2831.html.

WHO (2003): World Health Report 2003. World Health Organization 2003.

WHO (2008): World Malaria Report 2008. World Health Organization 2008.

YAMEY, G. (2004): Roll Back Malaria: a failing global health campaign, *BMJ*, 328(7448), pp. 1086–1087.

Appendix

A Identification

We are interested in the average effect of the intervention at time $t = 1$. We first rewrite the parameter of interest in terms of the populations that do not change behavior because of the intervention ($B_1^0 = B_1^1 = 1$ or $B_1^0 = B_1^1 = 0$), that change behavior in the desired way (compliers, $B_1^0 = 0, B_1^1 = 1$) and that change behavior in the opposite way (defiers, $B_1^0 = 1, B_1^1 = 0$). We then exploit that the effect for the first population is zero (A1), and that the probability of observing a defier is zero (A2).

$$
\begin{aligned}
E(Y_1^1 - Y_1^0) &= E(Y_1^1 - Y_1^0 | B_1^0 = 0, B_1^1 = 0) P(B_1^0 = 0, B_1^1 = 0) \\
&\quad + E(Y_1^1 - Y_1^0 | B_1^0 = 1, B_1^1 = 1) P(B_1^0 = 1, B_1^1 = 1) \\
&\quad + E(Y_1^1 - Y_1^0 | B_1^0 = 0, B_1^1 = 1) P(B_1^0 = 0, B_1^1 = 1) \\
&\quad + E(Y_1^1 - Y_1^0 | B_1^0 = 1, B_1^1 = 0) P(B_1^0 = 1, B_1^1 = 0) \\
&\overset{A1}{=} E(Y_1^1 - Y_1^0 | B_1^0 = 0, B_1^1 = 1) P(B_1^0 = 0, B_1^1 = 1) \\
&\quad + E(Y_1^1 - Y_1^0 | B_1^0 = 1, B_1^1 = 0) P(B_1^0 = 1, B_1^1 = 0) \\
&\overset{A2}{=} E(Y_1^1 - Y_1^0 | B_1^0 = 0, B_1^1 = 1) P(B_1^0 = 0, B_1^1 = 1)
\end{aligned}
$$

Now consider $E(Y_1^1 - Y_1^0 | B_1^0 = 0, B_1^1 = 1)$. Let $X_1 \equiv X_1^B \cap X_1^Y$. We first rewrite the parameter of interest in terms of the behavior-specific potential outcomes. We apply the definition of a complier and exploit that the intervention has no effect if behavior is unchanged (A1). Next we use the definition of conditional expectations and apply Bayes' rule. We then exploit the conditional independence assumption (A4) to rewrite the effect of interest in terms of observable objects.

$$
\begin{aligned}
&E(Y_1^1 - Y_1^0 | B_1^0 = 0, B_1^1 = 1) \\
&= E[B_1^1 Y_1^{11} + (1 - B_1^1) Y_1^{10} - B_1^0 Y_1^{01} - (1 - B_1^0) Y_1^{00} | B_1^0 = 0, B_1^1 = 1] \\
&= E(Y_1^{11} - Y_1^{00} | B_1^0 = 0, B_1^1 = 1) \\
&\overset{A1}{=} E(Y_1^{11} - Y_1^{10} | B_1^0 = 0, B_1^1 = 1) \\
&= \int E(Y_1^{11} - Y_1^{10} | X_1 = x) f_{X_1 | B_1^0 = 0, B_1^1 = 1}(x) dx \\
&= \frac{\int E(Y_1^{11} - Y_1^{10} | X_1 = x) P(B_1^0 = 0, B_1^1 = 1 | X_1 = x) f_{X_1}(x) dx}{P(B_1^0 = 0, B_1^1 = 1)} \\
&\overset{A4}{=} \frac{\int [E(Y_1^{11} | B_1 = 1, X_1 = x) - E(Y_1^{10} | B_1 = 0, X_1 = x)] P(B_1^0 = 0, B_1^1 = 1 | X_1 = x) f_{X_1}(x) dx}{P(B_1^0 = 0, B_1^1 = 1)} \\
&= \frac{\int [E(Y_1 | B_1 = 1, X_1 = x) - E(Y_1 | B_1 = 0, X_1 = x)] P(B_1^0 = 0, B_1^1 = 1 | X_1 = x) f_{X_1}(x) dx}{P(B_1^0 = 0, B_1^1 = 1)}
\end{aligned}
$$

Consequently, we obtain for the parameter of interest

$$
\begin{aligned}
&E(Y_1^1 - Y_1^0) \\
&= \int [E(Y_1 | B_1 = 1, X_1 = x) - E(Y_1 | B_1 = 0, X_1 = x)] P(B_1^0 = 0, B_1^1 = 1 | X_1 = x) f_{X_1}(x) dx
\end{aligned}
$$

Now consider $P(B_1^0 = 0, B_1^1 = 1 | X_1 = x)$. If there are no defiers, the probability of being a complier is equal to the difference in the probability to show the desired behavior with and without the intervention. If we can then proxy no-intervention behavior with pre-intervention behavior, this is

identified from observed objects.

$$P(B_1^0 = 0, B_1^1 = 1 | X_1 = x)$$
$$\stackrel{A2}{=} [P(B_1^0 = 0, B_1^1 = 1 | X_1 = x) + P(B_1^0 = 1, B_1^1 = 1 | X_1 = x)]$$
$$\quad - [P(B_1^0 = 1, B_1^1 = 1 | X_1 = x) + P(B_1^0 = 1, B_1^1 = 0 | X_1 = x)]$$
$$= P(B_1^1 = 1 | X_1 = x) - P(B_1^0 = 1 | X_1 = x)$$
$$= E(B_1^1 - B_1^0 | X_1 = x)$$
$$\stackrel{A3}{=} E(B_1^1 - B_0^0 | X_1 = x)$$
$$= E(B_1 | X_1 = x, I = 1) - E(B_0 | X_1 = x, I = 0)$$

The last equality follows from everyone being treated. Consequently, we obtain for the parameter of interest

$$E(Y_1^1 - Y_1^0)$$
$$= \int E(Y_1^{11} - Y_1^{10} | X_1 = x) E(B_1^1 - B_1^0 | X_1 = x) f_{X_1}(x) dx$$
$$= \int [E(Y_1 | B_1 = 1, X_1 = x, I = 1) - E(Y_1 | B_1 = 0, X_1 = x, I = 1)]$$
$$\quad \times [E(B_1 | X_1 = x, I = 1) - E(B_0 | X_1 = x, I = 0)] f_{X_1}(x | I = 1) dx$$

where the last line only depends on observed objects and where all elements are non-parametrically identified without additional assumptions.

B Asymptotic distribution of the estimator

Here we present the asymptotic distribution of our estimator as implemented in the application. We use consistent and asymptotically normally distributed parametric estimators for the probabilities that appear in the weights of the IPW estimator of the parameter of interest. These estimators can be interpreted as GMM estimators where the moment conditions correspond to the derivatives of the (log-)likelihood function of the maximum-likelihood estimators with respect to the parameter vector of interest:

$$\begin{pmatrix} g_0(X_0, B_0, \beta_0) \\ g_1(X_1, B_1, \beta_1) \end{pmatrix} = \begin{pmatrix} \partial L_0(\cdot, \beta_0)/\partial \beta_0 \\ \partial L_1(\cdot, \beta_1)/\partial \beta_1 \end{pmatrix} \tag{13}$$

where the subscript 0 refers to the estimation of $p_0(X_0, \beta_0) \equiv P(B_0 = 1 | X_0 = x)$, and where the subscript 1 refers to the estimation of $p_1(X_1, \beta_1) \equiv P(B_1 = 1 | X_1 = x)$.

One complication that arises in our specific case is that $p_0(X_0, \beta_0)$ is estimated on the data from $t = 0$ while $p_1(X_1, \beta_1)$ is estimated on the data from $t = 1$. However, the approach proposed by Newey (1984) can still be applied with the following adaptation. Assume that the data from $t = 0$ and $t = 1$ are stacked on top of each other resulting in a sample size of $N = N_0 + N_1$ and variables $X = (X_0 \ X_1)'$, $B = (B_0 \ B_1)'$ and $Y = (Y_0 \ Y_1)'$. Moreover, introduce an indicator variable D that is equal to one if $t = 1$ and zero otherwise. If only observations from $t = 1$ should be used to estimate a given set of parameters, all variables are multiplied by D. If only observations from $t = 0$ should be used, all variables are multiplied by $1 - D$. Furthermore, the number of observations has to be corrected in case averages are used. Thus, the moment conditions that use the stacked data are

$$g(D, X, B, \beta) \equiv \begin{pmatrix} g_0(D, X, B, \beta_0) \\ g_1(D, X, B, \beta_1) \end{pmatrix} = \begin{pmatrix} \partial L_0(\cdot, \beta_0)/\partial \beta_0 \\ \partial L_1(\cdot, \beta_1)/\partial \beta_1 \end{pmatrix}. \tag{14}$$

In the GMM framework, the estimators of the unknown parameter vector $\beta = (\beta_0 \ \beta_1)'$ satisfy the

conditions

$$g(d_i, x_i, b_i; \hat{\beta}) \equiv \begin{pmatrix} \frac{1}{N}\sum_{i=1}^{N} \frac{N}{N_0} g_0(d_i, x_i, b_i; \hat{\beta}_0) \\ \frac{1}{N}\sum_{i=1}^{N} \frac{N}{N_1} g_1(d_i, x_i, b_i; \hat{\beta}_1) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{15}$$

The estimator we propose can be interpreted as a two-step GMM estimator with the first step given above. The second step is an estimator of the form

$$\hat{\theta} = \frac{1}{N}\sum_{i=1}^{N} \frac{N}{N_1} w_i(d_i, x_i, b_i; \hat{\beta}) y_i \tag{16}$$

where

$$w_i(d_i, x_i, b_i; \hat{\beta}) \equiv \left[\frac{b_i}{p_1(x_i; \hat{\beta}_1)} - \frac{1-b_i}{1-p_1(x_i; \hat{\beta}_1)} \right] \left[p_1(x_i; \hat{\beta}_1) - p_0(x_i; \hat{\beta}_0) \right] d_i. \tag{17}$$

We can write the corresponding GMM estimator as

$$\frac{1}{N}\sum_{i=1}^{N} \frac{N}{N_1} h(d_i, x_i, b_i; \hat{\beta}, \hat{\theta}) = 0 \quad \text{where} \quad h(d_i, x_i, b_i; \hat{\beta}, \hat{\theta}) \equiv \hat{\theta} - w_i(d_i, x_i, b_i; \hat{\beta}) y_i. \tag{18}$$

Newey (1984) shows that for the special case where the first-step moment conditions for β do not depend on the second-step parameter θ, as is the case here, this type of estimator is consistent and asymptotically normally distributed if a set of regularity conditions are satisfied (Newey, 1984, p. 202). The conditions include stationary and ergodic data, finite parameters that are not at the boundary of the parameter space, existence, measurability and continuity of the first and second-moment functions and their derivatives, as well as full rank of the derivatives of the moment conditions with respect to the parameters. Furthermore, the sample moments should converge to their population counterparts with diminishing variance (consistent estimators) and should yield uniquely identified solutions for the unknown parameters. These conditions are considered rather weak, especially with independent and identically distributed data in the dimension i as in random samples of repeated cross sections (see also Hansen, 1982; Newey and McFadden, 1994). Applying the result by Newey (1984) our estimator is consistent and asymptotically normally distributed with asymptotic covariance matrix

$$AVar(\sqrt{N}\hat{\theta}) = H_\theta^{-1}\left(V_{hh} + H_\beta G_\beta^{-1} V_{gg}(G_\beta^{-1})' H_\beta' - H_\beta G_\beta^{-1} V_{gh} - V_{hg}(G_\beta^{-1})' H_\beta'\right)(H_\theta^{-1})' \tag{19}$$

where

$$H_\theta \equiv E\left[\frac{\partial h(\cdot)}{\partial \theta}\right] = 1 \tag{20}$$

$$H_\beta \equiv E\left[\frac{\partial h(\cdot)}{\partial \beta}\right] = -E\left[\frac{\partial w_i(d_i, x_i, b_i, \beta)}{\partial \beta} y_i\right] \tag{21}$$

$$G_\beta \equiv E\left[\frac{\partial g(\cdot)}{\partial \beta}\right] \tag{22}$$

$$V_{hh} \equiv E[h(\cdot)^2] = Var(w_i(d_i, x_i, b_i, \beta) y_i) \tag{23}$$

$$V_{gg} \equiv E[g(\cdot)g(\cdot)'] \tag{24}$$

$$V_{gh} \equiv E[g(\cdot)h(\cdot)]; \quad V_{hg} \equiv V_{gh}'. \tag{25}$$

Consequently, we obtain the following specific asymptotic covariance matrix for our estimator:

$$
\begin{aligned}
AVar(\sqrt{N}\hat{\theta}) \;=\; & Var(w_i(d_i, x_i, b_i, \beta)y_i) \\
& + E\left[\frac{\partial w_i(\cdot)}{\partial \beta}y_i\right] G_\beta^{-1} V_{gg}(G_\beta^{-1})' E\left[\frac{\partial w_i(\cdot)}{\partial \beta}y_i\right]' \\
& + E\left[\frac{\partial w_i(\cdot)}{\partial \beta}y_i\right] G_\beta^{-1} V_{gh} + V_{hg}(G_\beta^{-1})' E\left[\frac{\partial w_i(\cdot)}{\partial \beta}y_i\right]'
\end{aligned}
\tag{26}
$$

Note that if the first-step parameters β were known and not estimated, the variance formula would simplify to $AVar(\sqrt{N}\hat{\theta}) = Var(w_i(d_i, x_i, b_i, \beta)y_i)$.

C Descriptive statistics for our sample

Table 4: Descriptive statistics for our sample

	2000		2004	
	Mean	Std. Dev	Mean	Std. Dev
Child characterstics				
Single birth	0.96	0.20	0.97	0.16
Age	0.51	0.31	0.51	0.30
Female	0.50	0.50	0.49	0.50
Size at birth				
Very large	0.23	0.42	0.31	0.46
Average	0.59	0.49	0.49	0.50
Very small	0.17	0.38	0.16	0.37
Place of birth				
Home, other	0.48	0.50	0.32	0.47
Public hospital	0.38	0.49	0.38	0.49
Private hospital	0.14	0.35	0.30	0.46
Mother characteristics				
Age (years)	26.14	6.85	26.09	6.54
16-19 years	0.15	0.36	0.13	0.34
20-24 years	0.33	0.47	0.36	0.48
25-29 years	0.24	0.43	0.24	0.43
30-34 years	0.13	0.33	0.14	0.35
35-39 years	0.09	0.29	0.09	0.28
40-44 years	0.04	0.18	0.04	0.18
45-49 years	0.02	0.12	0.01	0.09
Highest education				
No education	0.28	0.45	0.24	0.43
Primary	0.64	0.48	0.65	0.48
Secondary, higher	0.08	0.27	0.10	0.30
Ethnicity				
Chewa	0.29	0.45	0.36	0.48
Tumbuka	0.10	0.30	0.10	0.30
Lomwe	0.19	0.39	0.17	0.38
Yao	0.15	0.36	0.16	0.37
Ngoni	0.09	0.29	0.08	0.27
Other	0.19	0.39	0.13	0.34
Religion				
Christian	0.82	0.39	0.82	0.38
Muslim	0.17	0.37	0.17	0.37
None, other	0.02	0.12	0.01	0.12
Number of children ever born				
1	0.23	0.42	0.21	0.41
2	0.21	0.40	0.20	0.40
3	0.16	0.36	0.17	0.37
4	0.12	0.32	0.13	0.34
5+	0.29	0.46	0.29	0.45
Knowledge of oral rehydration				

Table 4: Descriptive statistics for our sample (*continued*)

	2000		2004	
	Mean	Std. Dev	Mean	Std. Dev
Never heard of	0.15	0.36	0.07	0.25
Heard of	0.12	0.33	0.18	0.39
Uses	0.72	0.45	0.75	0.44
Married	0.89	0.31	0.88	0.32
Labor market status				
Not working	0.41	0.49	0.39	0.49
Worked in the past year	0.05	0.21	0.04	0.19
Currently working	0.55	0.50	0.57	0.49
Problems to get medical help				
Where to go big problem	0.17	0.37	0.16	0.36
Permission to go big problem	0.08	0.27	0.08	0.28
Money big problem	0.55	0.50	0.35	0.48
Taking transport big problem	0.53	0.50	0.59	0.49
Not wanting to go alone big problem	0.24	0.43	0.27	0.45
Concern no female provider big problem	0.13	0.34	0.15	0.36
Visited by a family worker in past 12 months	0.14	0.35	0.12	0.32
Any older children died	0.08	0.28	0.08	0.27
Mother's height (cm)	1540	164	1484	331
Height: <1.50 m	0.16	0.37	0.16	0.37
Height: 1.50-1.59 m	0.61	0.49	0.58	0.49
Height: 1.60-1.69 m	0.21	0.41	0.20	0.40
Height: >1.70 m	0.01	0.09	0.01	0.10
Regional characterstics				
Altitude (m)	822	333	864	339
<250 m	0.04	0.19	0.04	0.19
250-499 m	0.46	0.50	0.41	0.49
500-749 m	0.14	0.34	0.15	0.36
750-999 m	0.26	0.44	0.28	0.45
1000-1249 m	0.08	0.27	0.10	0.30
1250+ m	0.02	0.15	0.02	0.14
Distance to lake (km)	52.72	38.74	54.67	38.10
<50 km	0.52	0.50	0.49	0.50
50-99 km	0.36	0.48	0.38	0.49
100-149 km	0.10	0.30	0.11	0.31
150+ km	0.01	0.12	0.02	0.14
Region				
Northern	0.17	0.38	0.13	0.34
Central	0.37	0.48	0.41	0.49
Southern	0.46	0.50	0.46	0.50
Rural area	0.84	0.36	0.93	0.26
Distance to next city (km)	18.85	12.73	21.69	13.52
0-19 km	0.59	0.49	0.52	0.50
20-39 km	0.35	0.48	0.37	0.48
40-60 km	0.06	0.25	0.10	0.30
60+ km	0.00	0.00	0.01	0.08
Partner characteristics				
Has no partner	0.11	0.31	0.12	0.33
Partners age (years)	28.67	13.08	27.98	13.15
<20 years	0.01	0.08	0.01	0.09
20-29 years	0.39	0.49	0.40	0.49
30-39 years	0.32	0.47	0.31	0.46
40-50 years	0.13	0.33	0.13	0.33
60+ years	0.04	0.20	0.03	0.18
Highest education				
No education	0.13	0.33	0.13	0.34
Primary	0.59	0.49	0.56	0.50
Secondary, higher	0.17	0.37	0.19	0.39
Household characterstics				
Wealth index	-0.12	0.74	-0.16	0.69
Poorest quintile	0.20	0.40	0.19	0.39
Poorer quintile	0.23	0.42	0.27	0.44

	2000		2004	
	Mean	Std. Dev	Mean	Std. Dev
Middle quintile	0.19	0.39	0.26	0.44
Richer quintile	0.21	0.41	0.15	0.35
Richest quintile	0.16	0.37	0.14	0.34
Number of observations	2770		2523	

Note: If fractions in subcategories do not add up to 1 the remainder is N/A.

D Complete Probit estimation results

Table 5: Complete Probit estimation results

	Pre-intervention (2000)			Post-intervention (2004)		
	coeff.		t-stat.	coeff.		t-stat.
Constant	-0.4687		-0.56	0.9579	*	1.66
Child characteristics						
Female	0.0421		0.49	0.0319		0.50
Age (years)	0.1160		0.78	0.0398		0.38
Size at birth (reference: average, N/A)						
Very large	0.0276		0.28	-0.0433		-0.61
Very small	0.0270		0.23	-0.1166		-1.28
Place at birth (reference: home, other, N/A)						
Public hospital	0.1050		1.10	0.3429	***	4.42
Private hospital	0.0477		0.38	0.2858	***	3.48
Mother characteristics						
Age	-0.0099		-0.39	-0.0114		-0.58
Age categories (reference: age 20-24)						
<20	-0.2770		-1.64	0.0321		0.23
25-29	0.2594		1.46	0.2494	*	1.85
30-34	0.3565		1.25	0.2943		1.39
35+	-0.0038		-0.01	0.1562		0.48
Highest education (reference: primary)						
No education	-0.4323	***	-4.13	-0.2448	***	-2.93
Secondary, higher	0.4804	***	3.05	0.3806	***	2.70
Ethnicity (reference: Chewa, N/A)						
Tumbuka	-0.0642		-0.32	-0.1143		-0.73
Lomwe	-0.4373	**	-2.27	-0.1581		-1.20
Yao	0.1645		0.89	0.3453	**	2.23
Ngoni	-0.3034	*	-1.70	-0.0575		-0.42
Other	0.0796		0.47	-0.1738		-1.29
Religion: Christian	0.2094		1.44	0.1024		0.85
Number of children ever born (reference: 1)						
2	-0.0449		-0.34	0.3561	***	2.90
3	-0.2593		-1.55	0.1172		0.85
4	-0.3749	*	-1.85	0.2432		1.52
5+	-0.2525		-1.26	0.1503		0.88
Knowledge of oral rehydration (reference: not known, N/A)						
Heard of	0.1838		1.38	0.0101		0.07
Used	0.2498	*	1.85	0.0185		0.22
Labor market status: not working	0.0115		0.13	-0.0265		-0.38
Problems to get medical help (reference: no problem, N/A)						
Where to go	-0.1544		-1.05	-0.1367		-1.45
Permission to go	-0.0968		-0.56	0.0423		0.31
Money	0.0149		0.18	-0.0831		-1.08
Taking transport	0.0386		0.46	-0.1612	**	-2.24
Not wanting to go alone	-0.0011		-0.01	0.1372	*	1.72
Concern no female provider	0.1824		1.37	-0.1736	*	-1.85
Visited by a family worker in past 12 months	-0.0301		-0.27	0.0493		0.49
Any older children died	0.1176		0.88	0.1567		1.37

Table 5: Complete Probit estimation results (*continued*)

	Pre-intervention (2000)			Post-intervention (2004)		
	coeff.		t-stat.	coeff.		t-stat.
Mother's height (cm)	-0.0002		-0.62	0.0001		1.37
Mother's height category (reference: height 1.50-1.59 m, N/A)						
Height: <1.50 m	-0.0632		-0.55	-0.1461	*	-1.65
Height: >1.60 m	-0.1176		-1.10	0.0085		0.10
Regional characteristics						
Altitude (m)	-0.0015	***	-4.26	-0.0017	***	-6.32
Altitude (reference: <750 m)						
750-1000	-0.0518		-0.26	0.1189		0.91
1000+	0.2340		0.88	0.6533	***	2.97
Distance to lake (km)	-0.0100	***	-2.71	-0.0067	***	-2.80
Distance to lake (reference: distance <50 km)						
50-99	0.1175		0.62	0.0408		0.32
100+	0.1121		0.29	0.6160	**	2.28
Region (reference: Southern)						
Northern	0.8275	***	4.51	0.3492	**	2.24
Central	0.3212	**	2.25	0.1866		1.60
Urban area	0.0575		0.39	-0.4508	***	-3.11
Distance to next city	0.0135		1.60	0.0010		0.17
Distance to next city (reference: distance 0-19)						
20-39	-0.1666		-0.89	-0.0976		-0.80
40+	0.1246		0.39	-0.1740		-0.70
Partner characteristics						
Has no partner	0.7657	*	1.94	-0.1382		-0.41
Highest education (reference: primary, N/A)						
No education	-0.1381		-1.04	-0.1177		-1.12
Secondary, higher	0.5236	***	4.61	0.2456	***	2.58
Partners age	0.0212	*	1.95	-0.0004		-0.04
Partners age (reference: age <30-39, N/A)						
<25	0.1408		0.75	0.1685		1.03
25-29	-0.0245		-0.18	0.0316		0.27
40+	0.1508		0.75	0.1108		0.72
Household characteristics						
Wealth index	0.5937	***	6.37	0.4566	**	2.55
Wealth index (reference: poorest quintile, N/A)						
Poorer quintile	-0.1501		-1.10	-0.0218		-0.22
Middle quintile	0.3468	***	2.60	0.3474	***	3.44
Richer quintile	0.0324		0.25	0.0383		0.32
Richest quintile	0.1281		0.62	0.3106		1.29

Note: ***/**/* indicates significance on the 1/5/10% level.

E Loss of observations

Table 6: Loss of observations

	Included	Excluded	Excluded %
Without support enforcement	2523	0	0
Support within largest nontreated and smallest treated $\hat{p}_1(x_{1i})$	2450	73	2.89
Support within 10th largest nontreated and 10th smallest treated $\hat{p}_1(x_{1i})$	2295	228	9.04
Trim weights to a maximum of 1/0.1*	2450	73	2.89
Trim weights to a maximum of 1/0.05*	2450	73	2.89
Trim weights to a maximum of 1/0.01*	2450	73	2.89
Discard observations with weight > 1/0.1*	2439	84	3.33
Discard observations with weight > 1/0.05*	2449	74	2.93
Discard observations with weight > 1/0.01*	2450	73	2.89

Note: *73 observations are lost due to enforcement of support within largest nontreated and smallest treated $\hat{p}_1(x_{1i})$.