STOCKHOLM SCHOOL OF ECONOMICS
Department of Economics
5350 Master's thesis in economics
Academic year 2014–2015

Does Health Aid Reduce Infant Mortality in Malawi?

Mårten Ambjörnsson (40600) and Simon Costa (40601)

In this paper we analyze whether aid projects directed to the health sector significantly reduce infant mortality in Malawi. This adds to the aid effectiveness debate, where most works are cross-country studies, as opposed to our subnational approach. We use geocoded data from the Domestic Health Survey 2010 (DHS) and from the AidData initiative, which covers 80 % of aid donations to Malawi from 2002. We analyze the years 2002 to 2008 and use fixed effects and difference in differences models on district level. We also analyze infant mortality on local level, looking at aid projects in proximity of observed births. On local level we aim to overcome the selection problem of non-random aid allocation with an instrumental variable and matching. We cannot reject the hypothesis of no effect of aid to the health sector on infant mortality. We find that aid projects are not allocated randomly nor solely according to need.

Keywords: Aid effectiveness, Infant Mortality, Health, Development

JEL: I120, I130, O150, O220

Supervisor: Maria Perrotta Berlin
Date submitted: May 14, 2015
Date examined: May 28, 2015
Discussants: Elise Lagerborg, Rokas Narkus
Examiner: Örjan Sjöberg

We would like to express our gratitude to our supervisor Maria Perrotta Berlin for inspiration, useful remarks and engagement throughout the writing process of this master thesis. We would also like to thank the Stockholm School of Economics as well as our fellow students for many discussions and support throughout this endeavor.

Contents

List of tables

1. Introduction

One central question in development economics is aid effectiveness. This study adds to this field by studying the impact of health aid projects on infant mortality at a subnational level in Malawi. Our research question is, whether infant mortality has decreased in the presence of aid, where the null hypothesis is no effect of aid. Geocoded aid data and Demographic Health Survey data enable us to test this hypothesis by estimating the effect on district level, as well as local level. We use a panel model, difference in difference estimation, an instrumental variable approach and matching to estimate the effect of health aid projects on infant mortality. While control variables show the expected impact, we cannot reject the null hypothesis of no effect of health aid on infant mortality.

Effectiveness and allocation of aid have generated a large literature on cross-country level, such as Burnside and Dollar (2000), who ask whether aid had a positive impact on growth of developing countries. Subnational studies on the impact of aid are, however, rare (De, Becker 2015). New data on subnational level gives the opportunity to address these questions anew and reduce problems with cross-country studies. The focus of most cross-country studies lies on the impact of aid on economic growth, while the impact of aid on other outcomes than economic progress, such as health, is less researched. This is problematic for a number of reasons, some of which are summarized by Bourguignon and Sundberg (2007). In short, they find that not all aid projects are directly aimed at economic outcomes and that they will not affect growth immediately. Furthermore different donors might follow different agendas and it is statistically complicated to isolate the effect of aid on growth. Conflict and different levels of quality in the management of aid complicate the measurement of effects on cross-country level even further (Bourguignon, Sundberg 2007). De and Becker (2015) add that small, very effective projects might show no effect, because their effects are lost in noise when aggregate measures are used. They also point to the fact that GDP itself is often badly measured in aid-receiving countries.

All this calls for an analysis with more precision on subnational level. Geocoded aid data, i. e. data with GPS information on the location of projects, enable us to analyze the local effect of aid. The most extensive geocoded dataset on aid is available for Malawi, where eight percent of aid funds since 2002 are covered, which is why we chose this country. Malawi is among the least developed countries. In 2013 its GDP per capita was $780 and the Human Development Index ranks it on place 175. The population of 16.7 million people struggles with malnutrition, and famines are a recurring phenomenon. A high proportion of the population sustains itself from subsistence agriculture, and urbanization is low. The World Bank estimates the poverty rate for 2010 at 50.7 % (The World Bank 2015).

The aid data is furthermore categorized by sector. We can therefore look at aid directed at improving the health sector, which should allow for more precision than considering all aid. In order to measure the effectiveness of aid, we choose infant mortality as health outcome, because it is an indicator for the status of the overall health of a country (Greenstone, Hanna 2011). Reducing it is one of the Millennium development goals, not only because it is in itself a goal for any developed society, but also because its reduction leads to a more productive economy. Geocoded birth records available in the Demographic Health Survey (DHS) data from 2010 enable us to analyze the effect of health aid on the health outcome infant mortality, which is defined as the probability of infant death under the age of one year.

On district level, we calculate the number of health aid projects and disbursement and commitment of health aid per capita to estimate its impact on infant mortality. On individual level, we compute a variable of health aid presence. With this variable we regard all projects that should have a local impact and check for each observed birth, whether such projects are present within a ten- or twenty-kilometer radius. We then use several identification strategies to assess the impact of health aid presence on infant mortality. This specificity should circumvent the issue that a large share of aid is not directly targeted at GDP growth. Exploiting district and local variation should also decrease omitted variable bias in cross-country studies, by controlling for factors such as institutional differences (De, Becker 2015). Also the statistical problem of endogenous aid allocation might not be as prevalent within one country, or it might show different patterns than allocation according to need. Exact geographical information might give way to new instruments for aid allocation, if for example geographic accessibility plays a large role. Furthermore, aid effectiveness in each country is most likely different, which makes estimates specifically for Malawi interesting (Mishra, Newhouse 2007). Considering the goal of poverty alleviation, it is important to know to which extend aid can explain the progress made in improving health outcomes.

Summing up, this study is motivated by a number of reasons. The effect of health aid projects on infant mortality is interesting in itself as well as an overall indicator for the effect of health aid on the quality of health care in Malawi. The subnational level of this study is, furthermore, one step towards addressing the problems with cross-country studies in the assessment of the effectiveness of aid. Finally, geographic information should shed more light on the allocation of aid within Malawi.

In the following section we summarize the literature of the relevant research on aid effectiveness and infant mortality. We then motivate our empirical strategies and proceed with presenting our methods and results. Finally we discuss our results and conclude.

2. Literature Review

In this section we first give an overview over the effectiveness debate, to which we contribute. We then proceed with presenting studies on aid allocation, which are related to the identification problem in this work and then summarize studies on mortality and infant mortality to lay a foundation for the models we estimate.

2.1. Aid Effectiveness

Studies on the effectiveness of aid on an aggregate level focus on the impact on growth and remain controversial. While there is a trend for increased aid and policy makers do assume its effectiveness, research remains short of robust evidence on a macro level (Frot, Perrotta 2012). According to Frot and Perrotta (2012) as well as Bazzi and Clemens (2013) a major problem in estimating the effect of aid is selection bias: Aid is presumably allocated where it is needed the most. Instrumental strategies, which are used to overcome this problem, vary widely and lead to volatile results. Many of these strategies have been criticized and found to be invalid (Rajan, Subramanian 2008). The mixed evidence of studies finding positive effects, such as Burnside and Dollar (2000), and studies finding none, as Rajan and Subramanian (2008), leads to diverging interpretations. Easterly (2007) discusses arguments on why aid has failed to achieve higher growth, while others such as Collier (2006) argue that growth would have been much worse in the absence of aid. Bourguignon and Sundberg (2007) make the point that few observable effects are not surprising, once one considers aid in more detail. They argue that the heterogeneity of donors, incentives for policymakers and desired outcomes, might mean that purely economic improvement is hard to measure as an effect of aid. Considering that much aid is lost due to conflict and instability and that many projects might be poorly conceived and managed, together with statistical problems, they conclude that little or no visible effects on cross-country level come to no surprise.

On micro level the literature draws a different picture. In randomized control trials it has been shown, that different interventions can very well be effective. Examples of successful interventions aimed at health outcomes include Kremer and Miguel (2004) and Baird et al. (2011), who show that cheap deworming treatments in elementary schools have positive effects on school attendance and long-term labor supply. Similarly, Schultz (2004) evaluates the Mexican PROGRESA program, which entails conditional payments to women in eligible families. He also finds positive effects on school attendance, health, nutrition and a reduction in the severity of poverty. Another example is Svensson and Björkman (2009), who find community-based monitoring effective in increasing the quality of health care. Other studies look at infant mortality as a health outcome, in order to focus on other measures than GDP. To our knowledge, Boone (1996) is the first to look at infant mortality as an outcome of aid. He uses OECD data on Official Development Assistance from 1971 to 1990 and develops a model in which the government sector can decide whether to transfer aid flows or to use them to improve macroeconomic outcomes and HDI measures, such as infant mortality. He does not find significant effects of aid flows on either of these. Masud and Yontcheva (2005) follow to some extend Boone (1996). They focus on aid by NGOs as opposed to bilateral aid and use cross-country data from 1990 to 2001 that includes, apart from NGO aid and bilateral aid, several aggregate control variables which impact infant mortality, such as female literacy and urbanization. They find positive significant results and a larger effect of NGO aid than of bilateral aid. However, they also admit that their study suffers from several limitations – the endogeneity problem of aid allocation is not solved and their data on

NGO aid only includes aid co-financed by the European Commission. Mishra and Newhouse (2007) also look at the effect of health aid on infant mortality. They use a cross-country approach on data from 1970-2004 and several aggregate measures as controls. They find significant, but very small effects of health aid on infant mortality. They do not try to circumvent the problem of allocation and state that the estimated effects would likely be downwards biased, in case allocation is endogenous.

To our knowledge, De and Becker (2015) is the only study to date, which also attempts to assess aid effectiveness at a sub-national level and also addresses health aid with respect to health outcomes. Moreover the authors use the same geocoded dataset on Malawi together with a World Bank survey on living standards. They apply an instrumental variable approach and propensity score matching and find that health aid decreases disease severity, education aid increases school attendance and that water aid reduces diarrhea symptoms.

2.2. Allocation of Aid

Alesina and Dollar (2000) find that other factors than need impact where donations are allocated. By using a cross-country study and several indicators, they conclude that political considerations, differences in policies such as democratization, economic performance and colonial history between donors and recipients matter in the allocation of aid. Furthermore the allocation varies widely by donor. The Nordics for example are found to give more to countries with better institutions, while France is found to give most to countries with colonial ties. Lumsdaine (1993) presents evidence on donor spending and argues that humanitarian concerns are the main driver of foreign aid. Furthermore the fact that aid distribution across and within countries is insufficiently coordinated has drawn criticism (Rogerson, Steensen 2009). In this context the terms "aid darlings" and "aid orphans" are used to describe countries that receive too much or too little attention as recipients of aid. In case these patterns can be identified on a subnational level, this additional factor in the allocation of aid might provide an identification strategy for our setting, given that factors leading to the allocation, other than need, could be identified.

Due "to a lack of geocoded data almost no research on the allocation of aid on a sub-national level has been done. De and Becker (2015) also analyze geocoded aid data from Malawi and explain per capita expenditure of aid with data on living standards and regional variables. Using data on living standards implies that need is driving the allocation of aid, whereas using regional factors rests on the assumption that aid is also allocated with respect to feasibility, or what is known as "administrative convenience". Administrative convenience denotes factors such as good infrastructure and institutions, which facilitate aid projects and make it less costly to bring aid into place. To some extent these two factors work against each other, because relatively more developed areas also make it easier to implement aid projects. De and Becker (2015) estimate effects on per capita aid expenditure in three sectors: education, health and water. They find that aid allocation is non-random and that the three different sectors are strongly correlated with each other. The regions and living standard variables show different effects for the three sectors. According to their results, health aid is more allocated to areas with higher expenditure and more exposure to schooling, indicating an allocation to relatively more developed areas.

2.3. General Determinants of Mortality

Cutler et al. (2006) give an overview of research on the determinants of mortality in general, while they treat mortality in developing countries separately. The first result is that income increases life span and decreases child and infant mortality at all stages of development. Pritchett and Summers (1996) argue that income is the most important driver of health and all other factors will follow its development. Banister and Preston (1981) analyze improvements in China from the 1930s to the 1975 and suggest that a gain in life expectancy is likely caused by the combined effects of changes in income, literacy and the supply of calories. According to them one channel through which these factors have an impact on infant and child mortality is the improvement of health delivery, while calorie intake has a direct effect. Fogel (1997) analyzes declines in mortality since the early stages of the industrial revolution in Europe, with respect to a higher calorie intake. He finds that chronic malnutrition, and not famines, was a main driver of mortality. According to this study the reduction of chronic malnutrition led to vast declines in mortality in England and France from the 17th century onwards. However, interactions between public health development, higher productivity and incomes and calorie intake remain uncovered by this study.

Cutler et al. (2006) also point out that income facilitates the provision of public health infrastructure, i.e. sanitation and water. They also discuss one counter-argument, which follows from the observation of Cuba, where good public health measures increased health largely, while improvements in income were small at best.

Cutler et al. (2006) also take a closer look at the determinants of mortality in poor countries today and find that a much larger proportion of deaths is found among children living in those countries than in developed ones. Most of these deaths are caused by infectious diseases. They also observe that the vast improvements in India, China and Africa are attributable to an improvement in decreasing disease vectors, like anopheles mosquitoes and the immunization of children. Some diseases, which are leading causes of death, such as diarrheal disease and respiratory infections, are cheaply treatable. This makes health delivery and its quality key factors in decreasing infant mortality. The authors also note, that health status and health delivery vary greatly within countries. Socio-economic status determines much of the quality of care and, hence, income, race, education etc. should also be regarded in this respect.

2.4. Infant Mortality and Established Determinants

Infant and child mortality have been researched in many developing countries and several impacts on mortality rates are well established. One overview of variables which are available in DHS data and explain infant mortality is given in Mustafa and Odimegwu (2008). They group these into three categories, i. e. socio-economic, biological and demographic. They analyze DHS data from Kenya for 2003 and fit a logistic regression with the aim to rank the determinants of infant mortality. The most important determinant identified by their approach is breast feeding status followed by ethnicity, then fertility and the gender of the child. Lastly, maternal education and occupation also show significance.

More recently, environmental damage and its impact on health has also been addressed by Greenstone and Hanna (2011) where infant mortality is seen as an indicator for the overall state of the health care system. This also relates to the many deaths caused by cheaply treatable diseases in the developing world (Cutler, Deaton & Lleras-Muney 2006). Any determinant of health delivery will impact this measure, because health delivery is pivotal for the prevention of these deaths. Likewise, one can conclude that aid directed at improving the health sector should show improvements in infant mortality.

Another study using DHS data and infant mortality as a dependent variable is the one by Demombynes and Trommlerová (2012). They use data from 2003 and 2008 on Kenya, where infant and child mortality declined at the most rapid pace in Sub-Saharan Africa during this time period – at around 7-8 % per year. They estimate effects of the most commonly assumed determinants of child mortality, as well as the impact of living in different malaria risk zones across the country and the interaction of ownership of insecticide-treated bed nets (ITNs) and these risk zone factors. These are estimated with linear probability regressions. They then compare the effects found in 2003 and 2008 with a decomposition and find that ITNs have strong explanatory power, accounting for 58 % of the change, while most other factors – such as overall economic improvement, access to medication against the transmission of HIV from mothers to children, increased immunization – have no significant effect. A small effect can be found for improved sanitation. However, they admit that their method does not identify causal effects, but is more comparable to a measure of R-squared.

One factor reducing child and infant mortality is sanitation. In particular, Spears (2012) uses DHS data to examine the effect of India's Total Sanitation Campaign. He uses three different identification strategies: A panel with time and district fixed effects, difference in differences with parallel trend assumption on district level and a discontinuity approach stemming from discontinuous rewards for villages under the TSC. All three estimates find significant effects of latrines per capita on infant mortality on district level.

Summing up, infant mortality can be seen as a general measure for the status of health (Greenstone, Hanna 2011) and is accordingly impacted by most factors that impact public health, especially health delivery. Particularly important determinants found in the literature are income (Cutler, Deaton & Lleras-Muney 2006), sanitation (Spears 2012), maternal education and disease vectors (Mustafa, Odimegwu 2008), nutrition (Fogel 1997), as well as biological factors (Mustafa, Odimegwu 2008). It should also be mentioned that maternal education can be seen as an indicator of her overall socioeconomic status (Cutler, Deaton & Lleras-Muney 2006). At least some determinants, such as ethnicity and gender, depend to some degree on the specific country. This poses a difficulty in finding good controls for Malawi and makes the analysis of the most recent DHS data on Malawi relevant.

3. Data

In order to investigate the effect of health aid projects on infant mortality, we use a birth record of children born 2002 to 2008 from the demographic and health survey (DHS) in Malawi conducted in 2010, together with geocoded activity-level data from the government of Malawi's Aid Management Platform from AidData. One should note that the DHS data is only a sample of the population, whereas we have the whole population of aid projects. As the DHS 2010 on Malawi is sampled to be representative on district level it provides enough statistical power at our level of analysis. This section explains the geocoded aid data, followed by the DHS and the Malaria endimicity data. For clarity, more detailed descriptions of the data are presented in their respective context in the methods and result section.

3.1. Geocoded Aid Data

The dataset used is geocoded activity-level data from the government of Malawi's Aid Management Platform (Peratsakis et al. 2012). The data is made accessible by AidData, a partnership between the College of William & Mary, Development Gateway and Brigham Young University, with the purpose to track development funding. In total, the dataset covers 548 projects, at 2,523 locations and from thirty donor organizations. The projects in the dataset represent a total commitment of $5.3 billion. In comparison, the GDP of Malawi was $3.7 billion in 2013 (The World Bank 2015). The projects account for approximately eighty percent of the total external assistance reported to Malawi over the preriod 2000-2011.

This dataset is the first collection of sub-national geocoded locations of aid projects for any country. Projects in other countries have been geocoded after Malawi, but they do not cover such large shares of total external assistance or do not cover as many years.

Malawi, together with this dataset, is therefore very suitable for investigating the effectiveness of aid on a subnational level.

The geocoding is based on project documents gathered from donor offices in Malawi. The coding was performed using the AidData coding rules, see Strandow et al. (2011). This implies that the precision of the locations reported ranges from point locations, through two administrative regions (regions and districts in Malawi), to country level. In more detail, eight precision categories are assigned to the coordinates, so one can choose the level of precision needed for one's study. In this study, we are examining aid projects on districts in Malawi, using precision categories one to three, and on local level, using only exact point locations, i. e. using only precision category one.

The projects are divided into sixteen different sectors. The largest sectors are agriculture, education, health, integrated rural development and roads, public works and transportation. In this thesis we will look at the effect of projects in the health sector, as projects with a focus on health delivery can be assumed to be of particular importance for the effect on infant mortality. After examining the health projects we find that they, in general, have a focus on relatively broad health delivery improvements like basic health care, family planning and prevention of HIV. We, therefore, use all health aid projects and do not exclude any, as it seems plausible that they all should have an effect on overall health, of which infant mortality is an indicator.

Most projects in the dataset have information on year of agreement signed and year of planned completion. The range of agreement signed is from 1996 to 2015 and the range of planned completion is from 2002 to 2016. We use only projects that are signed before 2008, as we only have data on infant mortality till 2008. Some projects lack either year of agreement, year of planned completion or both. If the year of agreement is missing, we set it to the year of planned completion. If both dates are missing, the project is dropped.

There is also data on the cumulative commitment and disbursement in US Dollars for each project location. These numbers are only reported for whole projects, which often have multiple locations, and not separately for each location. We, therefore, divide total cumulative commitment and disbursement by the number of locations for each project. This approach assumes that commitment and disbursement are equally distributed between project locations. This is a relatively strong assumption, but it allows us to calculate an alternative measure of aid presence on district level. Cumulative commitment and disbursement for a given project differs in many cases, but aggregated to district level cumulative commitment and disbursement do not differ much.

3.2. DHS Data

To get data on births and infant survival for the studied time period, we use the latest standard Demographic and Health Survey (DHS) for Malawi. The survey was conducted in 2010 and carried out from June to November. Standard demographic and health surveys are national surveys conducted by the DHS Program funded by the United States Agency for International Development (USAID). The goal is to provide data for evaluation indicators in areas of population, health and nutrition in developing countries. The data was downloaded from the DHS Program website after obtaining permission (National Statistical Office (NSO) and ICF Macro 2015).

For sampling of the survey a two-stage cluster sampling procedure were used. At the first stage, on basis of enumeration areas taken from the 2008 Malawi Population and Housing Census, a predetermined number of clusters, which ensures representativeness at district level, were randomly selected. The probability of each enumeration area to be a selected cluster was proportional to its size. At the second stage, all households in a selected enumeration area were listed and selected for the questionnaire with equal probability. This procedure ensures representativeness of the sample for the population at district level. Less populated districts were oversampled, in order to take into account their smaller population. Urban areas were also oversampled, as most inhabitants in Malawi live in rural areas. In total, 849 clusters were selected from the census for the survey, 158 in urban areas and 691 in rural areas. This process provides the most detailed data about health indicators in Malawi available.

From the clusters, 27,307 households were selected in a systematic way. All selected households were visited. Eligible women (age span 15-49) were interviewed in all households and eligible men (age span 15-54) were interviewed in one third of the households. 25,311 of the selected households were occupied and 24,825 were successfully interviewed, giving a response rate of 98 %. In total 23,020 women (97 % response rate) and 7,175 men (92 % response rate) were interviewed. In order to generate data on fertility, all women were asked how many children they have given birth to during their life. For each child the women were asked about sex, date of birth, survival status and age at death for dead children. There is, therefore, a risk that the data on fertility might include errors as it is relying on the memory of the children's mothers. From the birth history it is then possible to get information on birth order, maternal age at birth, maternal age at first birth, year between births and multiple births.

The clusters in the survey are geo-referenced, giving us the location of the households. The coordinates are accurate to approximately fifteen to twenty meters, but to protect the integrity of the respondents the coordinates are randomly displaced. Urban clusters contain a minimum of zero and a maximum of two kilometers error. Rural clusters contain a minimum of zero and a maximum of five kilometers positional error, with a further one percent of the rural clusters displaced a minimum of zero and a maximum of ten kilometers. The clusters are, however, not displaced outside their respective districts. 23 clusters are not geocoded correctly as they have longitude and latitude equal to zero and, thus, these clusters are deleted.

In total, there are 72,301 observed births in the dataset born between 1973 and 2010. 70,659 remain after the incorrectly geocoded clusters are removed. Since there only is mostly data on aid presence from 2002 and onwards, all births before 2002 are dropped. All births after 2008 are also dropped, as children who have not lived a full year have not been fully exposed to mortality. There are then 27,071 births left in the sample. The location of households in the survey is the current location, a fact that may imply that it may not be the same location as where the children in the households were born. There is, however, information indicating whether a household has moved and all births observed before or in the same year as the one when the household moved are dropped, in order to ensure that all births are in the location stated in the survey. For the same reason, we also drop births by mothers classified as visitors.

We have then 21,784 observed births left to analyze. The births are relatively evenly distributed between the cohorts with the lowest number in 2002 with 2,371 births and the highest number in 2008 with 3,543 births, see the observed births column table A.I in appendix A. The distribution of the observed births is also relatively even between the districts, with an average of 806 births per district, see table A.II in appendix A.

3.3. Malaria Endemicity

Infectious diseases and particularly Malaria is a leading cause of death in Malawi, while malaria risk shows considerable variation in the country. We use data for Malaria endemicity as predicted mean parasite prevalence, mean PfPR, for children between two to ten years old in 2010, published by the Malaria Atlas Project (Gething et al. 2011). Because the estimates are for 2010, these risk rates are less precise for births before 2010. This risk measure is estimated via a Bayesian geospatial model based on Plasmodium Falciparum parasite rates (Gething et al. 2011) and published in a worldwide surface map. We follow Mustafa and Odimegwu (2008) and divide these risk rates into a categorical variable with low, high and very high malaria risk.

4. Empirical Strategies

In this section we give a motivation and an overview over the estimation strategies in this work. The main challenge with the estimation of the impacts of aid is non-random allocation. We attempt to solve this on district and local level.

On district level we analyze this question with two different models, using all aid projects coded on district level and lower. First, we aggregate the data to generate a panel over districts and years and apply fixed effects. This model yields unbiased estimates, if the treatment is assigned randomly, conditional on time and district fixed effects. Most of the factors that determine aid allocation could be considered to be time-invariant during our short time period of analysis. Examples of such factors are administrative convenience and geographical accessibility. Then the differences that lead to more or less aid presence are fixed and set in the beginning of the time period, this assumption does hold and a fixed effects estimator is unbiased. Second, instead of aggregating the observations, we use them as a repeated cross-section. We can then use a difference in difference approach, where the treatment happens on time and district level. The difference in difference estimator is unbiased, as long as the differences between treated and untreated are only in the initial levels of the outcome variable and the parallel trend assumption holds. In case the decision rule for the allocation of aid is only affected by time-invariant factors and the initial level of infant mortality, this first condition holds. Furthermore, the parallel trend assumption in this setting is, that infant mortality is assumed to change along the same trend in areas with and without aid projects. As the districts in Malawi are relative homogeneous, this assumption seems plausible.

We also analyze the impact of aid presence on local level, by looking at the number of health projects in proximity of observed births using only projects with exact point locations. The reason for doing this is that districts are relatively large areas and district level analysis might mask small-area patterns. We then also try to overcome the problem of reverse-causality, as explained below.

It is not clear in which way the outcomes targeted by aid, such as infant mortality, attract aid transfers. On a subnational level, aid might be allocated because of need. Other factors, such as administrative convenience and geographical accessibility, might also play an important role in the allocation process. In algebraic form, this reads as follows:

$$y_{i,s,t} = \gamma_s + \lambda_t + \beta \cdot aid_{i,s,t} + \delta \ \cdot X_{i,s,t,} + e_{i,s,t,}$$

Where y is the targeted outcome, X are controls and aid is the measure of aid presence. The indices are t for time, i for the individual and s is an index for the areas in which aid is allocated or not.

In case aid allocation is determined by variables Z, but also outcome y, the following relationship also has to be considered.

$$aid_{i,s,t} = \gamma_s + \lambda_t + \ \theta \cdot \ Z_{i,s,t} + \ \mu \cdot y_{i,s,t} + \ u_{i,s,t}$$

In this case reverse-causality is present and direct estimation of the first relationship leads to biased estimates.

One identification strategy to solve the problem with reverse-causality is instrumentation. The hope is to find geographic information, which could be used as an instrument. This instrument has to fulfill the assumptions of exclusion and relevance. That is, it has to be correlated with aid presence, while the outcome infant mortality is not directly impacted, but only via the allocation of aid. As most projects are located to district capitals and the degree of urbanization does not affect mortality rates, we use it as an instrument.

Matching estimators are also intuitive to use in this setting. The allocation of health aid projects can be seen as non-random allocation of a treatment. Matching methods have been developed, in order to estimate causal effects from observational data with nonrandomized treatment. They can be used, if available variables fulfill the assumptions of conditional independence of treatment status on the covariates and common support. We discuss these assumptions in more detail in the matching section. Intuitively, individuals that live in areas with aid presence can be matched to similar individuals in areas without the presence of health aid projects, in order to estimate the effect of health aid presence.

In the following methods and result section we explain the estimation strategies more detail and present the results for the models.

5. Methods and Results

This section first explains how the aid presence variable, the main independent variable of interest in this work, is constructed. The methods and results are then presented for the levels of analysis, i. e. on district and local level. For clarity, more detailed descriptions of the relevant data for each model are also presented under their respective section.

5.1. Aid Presence Variable

The main independent variable of interest in the following models is the aid presence variable. It is a measure of the number of active health projects in the area. The exact area depends on the model and is explained for each model in their respective section. This paragraph explains how the data from AidData is used to create this variable.

The first difficulty is that there is only information on the year of agreement signed and year of planned completion for each project. We do not know exactly when, after the agreement is signed, a project actually is in place and can be considered to be effective and have effect on infant mortality. We therefore use time-lags from zero to four years to allow for these different plausible time periods. More specifically, we calculate the number of projects in the area for each individual, signed before the year the individual was born (no lag). For one year time-lag we assume that it takes one more year for a project to become effective and only count projects signed at the latest two years before the birth. This calculation is repeated with lags up to four years. This means that the number of active projects declines when the number of lag increases. Once a project is counted as effective, it remains so for the remaining years. For the variable with a lag of two years, for example, a project signed in 2002 will be considered active for all years from 2005.

5.2. District Level

In this section we analyze the effect of aid presence on infant mortality on district level using all aid projects on district level and lower. The representativeness of the DHS on district level make it feasible to analyze the question on this level. It is done with two different models, an aggregated model and an individual model. The content of this section is as follows: first a summary of the data on district level is provided, followed by an examination of mortality rates in the districts. Finally, the two models are explained and their results presented.

5.2.1. Summary of Aid Presence Data

This section summarizes the aid presence data on district level. Malawi has 28 districts, but Likoma, a small group of islands in Lake Malawi, and Nkhata Bay districts are grouped together to one district, in order to fit with the DHS dataset. This give us a total of 27 districts. The districts are also divided into three larger regions: Southern, Central and Northern. For each district, we aggregate the three following variables for each year: the number of projects at district level or lower, cumulative commitment and disbursement. The districts vary a lot in population and size, so the cumulative commitment and disbursement are divided by the population of each district to get per capita values. The population is taken from the Malawi's National Statistics Office population and housing census for 2008 (Malawi National Statistical Office 2008). The population figures for 2008 are used for all years, as no yearly population figures are available for the districts. The aggregated values for 2002-2008 are shown in table A.II and figures A.I to A.III in appendix A.

There are in total 158 project locations on districts level for health projects signed before 2008. Approximately half of the locations are in the southern region. The number of projects per district varies a lot: Lilongwe, Mzimba and Zomba have all more than ten projects, while Neno has only one project. Total commitment is $109 million and total disbursement is $111 million, in all of Malawi, for the chosen projects. Approximately 70 % of the aid funding goes to the southern region, with the other two regions sharing the remaining aid relatively equally.

Average commitment is $8.4 and disbursement $8.5 per capita for the whole country. On a regional level the central region receives the least aid per capita, i. e. $3.5. The southern and northern regions receive more per capita than the average of Malawi, with the northern region receiving the highest figure, with approximately $13 per capita. The pattern is the same on district level. Per capita aid is highest in the southern and northern districts. It is lowest in districts in the central region and some districts in the southern region. The figures are, however, more diverse on district level, with Phlombe and Chiradzulu in the south receiving by far most aid funding, i. e. over $30 per capita, and Neno, Dowa, Mchinji receiving less than $2 per capita. Lilongwe, the national capital district, should also be mentioned, as its relatively high total number of projects does not translate to high commitment and disbursement per capita due to the large population in the district.

5.2.2. Mortality Rates

The mortality rates per district and year can be calculated by using a cohort life table approach, where the number of deaths at age twelve months or below for infants born during a year is divided by the number of total infants born that year. This is repeated for all districts and for all yearly cohorts. This procedure gives an estimate of the true probabilities of death and not only rates. The reason is that deaths of the whole cohort of one year are captured, as opposed to dividing the number of deaths in one year by births in one year. All children in the sample have been fully exposed to mortality. The probabilities are then multiplied by 1,000 to get the number of deaths per 1,000 children born. The average infant mortality rate for all children in the dataset for 2002-2008 is 72.6 deaths per 1,000 children born. The mortality rates for the different cohorts show a generally decreasing trend with the highest mortality for children born in 2002 and the lowest for children born in 2007, see table A.I in appendix A. The mortality per region varies a lot, indicating that it is interesting to look at our research question on a district level. At district level the mortality is lowest in the far north and in a couple of districts in the central region. Infant mortality is in these districts less than 60 deaths per 1000 children born. The mortality is highest in the southern region. Phalombe, in the south, has the highest mortality rate with 92.6 deaths per 1000 children born, see table A.II and map A.IV in appendix A.

The fact that infant mortality is highest in the south and lowest in the far north indicates that infant mortality is not a perfect determinant of health aid allocation, as aid funding per capita is highest in the northern region.

Table 1 shows the correlations between the aggregated variables of infant mortality, total number of projects, commitment and disbursement per capita on district level. The correlations are positive and range from 0.07 and 0.21.

Table 1 Correlation on district level between aid and infant mortality

Total number of aid projects	0.208
Commitment per capita	0.189
Disbursement per capita	0.072

5.2.3. Aggregated District Model

In this model we use the aggregated mortality and aid presence data at district level to create a panel for the districts in Malawi from 2002-2008. Panel data has many advantages, since it gives more observations and there are therefore more variability and more degrees of freedom in the regressions. We use a fixed effect OLS model, as the districts vary a lot in size, population and location etc. Each district has its own time-invariant intercept, which accounts for the heterogeneity between the districts. The advantage with the fixed effect model is that it controls for time-invariant confounding factors and, therefore, removes omitted variable bias that is fixed over time, and that we could not control for otherwise. A problem is that it is not possible to include time-invariant regressors, as the individual intercepts absorb all the heterogeneity. Infant mortality shows a decreasing trend over time, so the model is therefore extended with time fixed effects to allow the intercept to change over time as well. The general assumption of the fixed effect model is that treatment is randomly assigned, conditional on the observed covariates, time, and fixed effects (Wooldridge 2012).

In the aggregated model the observed births in the DHS data are aggregated to mortality rates for each district and year for 2002 to 2008, giving us a panel with 27 observations over seven years, i. e. in total 189 observations. The mortality rate is used as the dependent variable. As the independent variable of interest we use aid presence in the form of the number of active health projects in the district. We also apply two alternative specifications with commitment and disbursement per capita for active health projects respectively as the independent variable of interest.

Unfortunately, there is very little data on time-varying variables on district level in the DHS, as most data is only collected once at the time of the survey. It is therefore not possible to use controls in this model.

The estimated equation is:

$$mortality_{s,t} = \gamma_s + \lambda_t + \beta\, aid_{s,t} + e_{s,t}$$

Where aid is either the number of projects, commitment per capita or disbursement per capita in the district, depending on the specification. The indices are s for districts and t for time.

5.2.4. Results of the Aggregated District Model

The model is estimated for each of the three specifications of aid presence and for the no-year lag to four-year lag. The time fixed effects are all significant and the district fixed effects are jointly significant, indicating that a two-way fixed effect model is suitable. A regression for all estimates with the number of projects specification and no lag is presented in table B.I in appendix B. The estimated coefficients for the three specifications of aid presence and different lags are presented in table 2. As the estimated coefficients for the controls and the district and yearly fixed effects do not vary much between different specifications and lags, they are not presented for each lag and specification.

As the dependent variable is infant mortality in deaths per thousand children, the estimated coefficients for the specification with the number of projects can be interpreted, as the change in infant mortality when the number of aid projects in the district is increased by one. For the specification with commitment or disbursement per capita, the interpretation is the change in infant mortality for one more dollar in aid funding.

For the specification with the number of projects, all estimated coefficients are positive. The estimates are significant for no lag, one-year lag and three-year lag. For commitment and disbursement per capita the estimated coefficients for the no lag to two-year lag are negative and the rest are positive. The estimate of disbursement per capita for the two-year lag is the only significant estimate for the two aid funding specifications. The difference between the coefficients for commitment and disbursement is very small.

Table 2 Aggregated district model

Aid presence specification	Number of projects	Commitment per capita	Disbursement per capita
Aid presence, no lag	4.276**	-0.221	-0.166
	(1.643)	(0.350)	(0.293)
Aid presence, one-year lag	3.602*	-0.258	-0.360
	(2.035)	(0.358)	(0.242)
Aid presence, two-year lag	2.559	-0.421	-0.479*
	(2.553)	(0.387)	(0.284)
Aid presence, three-year lag	5.452*	0.284	0.361
	(3.234)	(0.546)	(0.654)
Aid presence, four-year lag	5.111	0.0674	0.0242
	(4.631)	(0.553)	(0.643)

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

5.2.5. Individual District Model

In the previous model all information is aggregated to district levels meaning, that all available individual information is lost, making it impossible to control for covariates at individual level. We cannot build a panel model with individuals, as that would require repeated observations of births, which we obviously do not have. Instead we have a repeated cross-section of births, which enables us to use a difference in difference approach, where it is sufficient that the treatment happens at time and on district level.

The treatment variable is the aid presence variable with differing treatment intensity over time and between districts.

This model assumes that the average change in infant mortality would have been the same for groups without treatment. For this assumption to hold it is, necessary that treatment is as good as randomly assigned given, controls and the time and district-specific effects, and that the parallel trend assumption is not violated. That is the trends in mortality rates between the districts should be the same in the absence of treatment. If a district is treated, it will induce a deviation from this trend. Also the level difference between districts should be captured by the district fixed effects (Angrist, Pischke 2008).

As the dependent variable is infant death, coded as zero if the infant survived its first year and as one if the infant died, the model is a linear probability model. The aid presence variable is on district level and its values are taken from the aggregate model.

The estimated equation is:

$$death_{i,s,t} = \gamma_s + \lambda_t + \beta \cdot aid_{s,t} + \delta \cdot X_{i,s,t} + e_{i,s,t}$$

Where aid presence is either the number of projects, commitment per capita or disbursement per capita in the district depending, on the specification and a vector of controls, X. The indices are i for individuals, s for districts and t for time.

The reason for adding controls is that unobserved factors might be correlated with aid presence and have an effect on infant mortality. In this case, the estimated coefficients of aid presence are biased. In order to account for this, we add control variables to the model. Since we are now using individual data, there are many possible controls available in the DHS data. As explained in the literature review, determinants of infant mortality are a mix of socio-economic, biological and demographical factors. One should note that the information is from the date the survey was carried out and not the date of birth. The wealth index is generated from the survey and it is determined by scoring households based on a set of characteristics, including access to electricity and ownership of various consumer goods. The households are then ranked and divided into quintiles. The following part explains which control variables are used and how they should affect infant mortality.

Socio-economic variables such as education, marriage status and maternal occupation, as well as household wealth are used as controls. Higher maternal education and more household wealth are supposed to decrease infant mortality, whereas the effect of occupation and marriage status is more ambiguous. Demographic variables such as religion and ethnicity may also be important, but their effect is also not very clear prior estimation. Maternal age at birth is necessary to include, as older females in general have more complications at childbirth and young females are often less experienced and in a more vulnerable position. The gender of the child is also controlled for, as females are more likely to survive as infants, which can vary between countries. Infectious diseases like HIV and malaria decrease the probability of survival for an infant. Controls for HIV-positive mother and malaria endemicity in the area are therefore added. The malaria risk factor is interacted with access to insecticide-treated bed nets to capture the effect of this preventive intervention. Fertility factors like birth order and birth interval are also an important determinant of mortality, as short birth intervals decrease the likelihood of survival. If the birth interval is not too short, higher order births should have a positive effect on incidence. Multiple births are often unexpected and mean a higher encumberment for the supporting household, increasing infant mortality. Another important determinant is the length of breast feeding, but unfortunately there are very few responses for this variable in the DHS dataset, so we choose to not include it to avoid reducing the sample size. Another important area that affects infant mortality is sanitary conditions, but these cannot be controlled for, as there is no suitable variable with enough observations.

When the model is estimated with all the controls described above some estimates are not significant. They are therefore dropped, as including them increase the variance of the estimated coefficients and since excluding the insignificant controls does not create any omitted variable bias. In the model presented below we include maternal education, maternal age at birth, sex of child, birth order, multiple births and HIV-positive mother. The HIV variable has missing values for some individuals and the sample is therefore restricted to 17,484 observations. Summary statistics of the control variables are provided in table A.III in appendix A.

5.2.6. Results of the Individual District Model

The model is estimated for each of the three specifications of aid presence and for no lag to four-year lag. The standard errors are clustered on district level, as standard OLS estimates are invalid under district level correlation. A regression with all estimated coefficients with the no lag in the aid presence variable and the number of projects specification is shown in table B.II in appendix B.

The dependent variable in the model is infant death as a binary variable, therefore the interpretation of the estimated coefficients is different compared to the aggregated district model. The estimated coefficients are interpreted as the change in probability of death when the variable is changed by one unit.

The signs of the control variables are in general as expected, indicating that our model is in line with previous research. Children whose mothers are between 20-35 years are significantly more likely to survive. Short birth interval, multiple births and a HIV-positive mother are all factors that significantly increase infant mortality. Female children are also more likely to survive their first year compared to males. One finding that is a bit surprising is that children to mothers with primary education are worse-off compared to children with mothers with no education. We cannot explain this, but one should, however, note that the estimate for secondary and higher education is largely negative even if it is not significant.

The estimated coefficients for the three specifications of aid presence and different lags are presented in table 3. As the estimated coefficients for the controls and the district and yearly fixed effects do not vary much between different specifications and lags, they are not presented for each lag and specification.

Table 3 Individual district model

Aid presence specification	Number of projects	Commitment per capita	Disbursement per capita
Aid presence, no lag	0.00285	-0.000221	-0.000131
	(0.00196)	(0.000162)	(0.000122)
Aid presence, one-year lag	0.00228	-0.000196	-0.000308
	(0.00246)	(0.000460)	(0.000283)
Aid presence, two-year lag	0.00128	-0.000384***	-0.000409***
	(0.00325)	(0.000136)	(0.000127)
Aid presence, three-year lag	0.00270	0.000146	0.000180
	(0.00259)	(0.000143)	(0.000203)
Aid presence, four-year lag	0.00461*	0.000233	0.000189
	(0.00230)	(0.000190)	(0.000371)

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

For the specification with the number of projects in the district, all estimated coefficients are positive and the estimate for the four-year lag is significant. For commitment and disbursement per capita the estimated coefficients are negative for no lag to two-year lag and positive for the rest. They are significant for the two-year lag. However, because the signs are not conclusive and the estimates do not draw a consistent picture, we believe that these effects are not robust. In case the model would yield conclusive results, one should also investigate the parallel trend assumption, which we do not deem meaningful under these circumstances.

The result of the two different district models is similar, indicating that the model choice is not that important. The result is, however, inconclusive as most estimates in the models are insignificant and the sign of the estimates differs between the specification with the number of projects per district and aid funding per capita.

5.3. Local Level

The result of the analysis on district level is inconclusive. One reason for this could be that the unit of analysis, i. e. districts, is relatively large area-wise. The analysis on district level might mask small-area patterns, as most aid projects do not have district-wide coverage. We, therefore, look at the number of health projects in proximity of a birth on local level using only projects with exact point locations. The main objective of the AidData geocoding is to record locations, where aid is actually committed or distributed. One can, therefore, argue that projects with exact point locations are active in their locations, i. e. have an effect on the health of the individuals in their proximity. If only the project administration and some functionalities are based in the location, but the actual coverage is larger it should be coded on district level instead.

For a project to be considered in proximity, we define that it must be within a certain radius of a birth. We choose the radiuses to be ten and twenty kilometers long, as we think they give a reasonable maximum distance for the individuals living in the area to have access to and be affected by the projects. The radius cannot be chosen much smaller than ten kilometers, as that would imply that very few births are observed close to projects and that the random displacement factor would have a very large effect. A larger radius than approximately twenty kilometers should not be chosen either, as the areas then get close to districts in size, and it is then better to do the analysis on district level. One should note that the clustering of the data together with the random displacement of the coordinates aggravate the analysis at local level, as the precise location of each household is not known. We can only calculate the distance between the cluster of the household and aid project locations. The problem is larger in rural areas, as the rural clusters are larger area-wise and the size of the displacement is generally larger. One would preferably like to have precise point locations of the households as well. This is, however, the data that is available and it should not pose too large problems, as the general idea of the local level approach is to get an indication of health aid presence in the area or community of the household and not analyzing exact distances.

This section continues as follows: first, we describe the projects with exact point locations and the number of infants born in proximity of these projects in more detail. Later, we explain how the model is set up and present its results. Finally, we discuss two approaches to improve upon this model: an instrumental variable approach and matching.

5.3.1. Projects with Exact Point Locations

There are 84 project locations with exact point locations for health projects signed until 2008, compared to 158 project locations on district level or lower. Of these 84 locations: 72 are located in 25 of the 27 district capitals and nine in four other cities. Only three projects are located outside the cities, all in the Phalombe district. In total, projects with exact locations are located in 32 geographical areas. These areas and their corresponding radiuses are shown in figures A.V and A.VI in appendix A.

In order to generate the aid presence variable with the number of projects in proximity, the geographical distances between the clusters in the DHS Malawi 2010 and all health aid projects are calculated. One should remember that the random displacement of the coordinates in the DHS implies that the maximum theoretical distance to a project is larger than the chosen radius.

2,964 births (13.6 %) are within a ten-kilometer radius of at least one health project, using the assumption of no lag. With a radius of twenty kilometers the number increases to 7,130 births. The share of children born close to projects does, however, change a lot per year. It is lowest in 2002 at three percent of the cohort and increasing to above twenty percent from 2006 for a ten-kilometer radius. The number of observed births is close to at least one project for all years, lags and radiuses are presented in table 4.

Table 4 The number of observed births close to at least one project for all years, lags and radiuses.

Ten-kilometer radius:

Year	No lag	One-year lag	Two-year lag	Three-year lag	Four-year lag
2002	75	7	-	-	-
2003	180	113	15	-	-
2004	196	196	112	13	-
2005	382	173	173	97	11
2006	674	375	186	186	93
2007	693	693	429	198	198
2008	764	764	764	459	227
Total	2,964	2,321	1,679	953	529

Twenty-kilometer radius:

Year	No lag	One-year lag	Two-year lag	Three-year lag	Four-year lag
2002	177	60	-	-	-
2003	436	258	78	-	-
2004	446	446	247	72	-
2005	995	404	404	223	78
2006	1,605	967	402	402	197
2007	1,669	1,669	1,067	442	442
2008	1,802	1,794	1,794	1,143	496
Total	7,130	5,598	3,992	2,282	1,213

As many projects are located in areas with more than one project, some observed births are in proximity of more than one project, especially in the later years. Of the 2,964 observed births within a ten-kilometer radius of at least one project, 1,438 are born close to more than one project if no lag is used, see table 5 for the distribution of the number of projects in proximity for both radiuses.

Table 5 Linear specification of aid presence

Number of projects	Ten-kilometer radius	Twenty-kilometer radius
0	18,820	14,654
1	1,526	2,537
2	829	2,293
3	260	1,133
4	63	412
5	134	142
6	93	188
7	37	103
8	22	207
9	-	115

5.3.2. Local Model

The model is a linear probability model with infant death as the dependent variable, coded as zero if the infant survived its first year and as one if the infant died. The independent variable of interest is aid presence, as the total number of aid projects in proximity in the year before the observed birth. The specification of the model is very similar to the individual district model explained previously. The dependent variable infant death, fixed effects for years and districts and the control variables are the same. The difference is the variable of interest. Instead of using aid presence at district level, aid presence in the proximity of the individuals is used. Dummy variables for years and districts are also added to control for yearly and district specific effects. The estimated equation is:

$$death_{i,s,t} = \gamma_s + \lambda_t + \beta \cdot aid_{i,s,t} + \delta \cdot X_{i,s,t} + e_{i,s,t}$$

Where aid is either the number of projects, commitment per capita or disbursement per capita in the district depending on the specification and a vector of controls, X. The indices are i for individuals, s for districts and t for time.

One could question the linear specification of the relationship between the number of aid projects and the effect on infant mortality used above, implying that for example two projects have twice the impact on infant mortality than one project. To get another measure of aid presence, the number of projects in proximity variable is transformed to a binary variable, coded as one if there is at least one project with exact point location in proximity the before the observed birth and zero otherwise, i. e. it just considers if there is a project or more in proximity and do not take into account the magnitude of the number of projects. The specification is called binary whereas the first specification is called linear.

The drawback of looking at projects on local level with exact point locations is that projects coded on district level are ignored. As they probably have an effect on infant mortality and the number of excluded projects coded on district level varies between districts, one should try to account for this in the aid presence variable. In order to do this we add the number of projects that are coded on district level with the linear aid presence variable, i. e. precise point locations in proximity and projects coded at the district level are assumed to have the same effect on infant mortality for the individual child. This specification is called "linear and district".

5.3.3. Results of the Local Model

A regression with all estimated coefficients with the no lag and the number of projects, as aid presence specification is shown in table B.III in appendix B. The estimated coefficients for the three specifications of aid presence, different lags and ten- and twenty-kilometer radiuses are presented in table 6. As the estimated coefficients for the controls and the district and yearly fixed effects do not vary much between different specifications and lags they are not presented for each lag and specification.

The estimated coefficients are interpreted as the change in probability of death when the aid presence variable is changed by one unit.

The estimated coefficients for the ten-kilometer radius are positive. The estimates for no lag and one year lag are significant for the linear and "linear and district" specification. The linear specification estimates have a larger magnitude compared to "linear district". No estimates are significant for the binary specification, but it has the largest variation in the estimated coefficients.

The result for the twenty-kilometer radius is a bit different. A majority of the estimated coefficients are negative and closer to zero, compared with the results for the ten-kilometer radius. All estimates for the linear specifications are negative, but only the estimate for the two-year lag is significant. The binary specification have both negative and positive estimated coefficients, no estimates are significant.

We also tried using a non-linear relationship of aid presence, by adding the square of the total number of projects together with the linear variable. It did not change the general result much compared to the linear specification and is, therefore, not presented.

In summary the result is inconclusive, as the ten- and twenty-kilometer radiuses have different signs of the estimated coefficients and few are significant.

Table 6 Local model, years 2002-2008

Ten-kilometer radius:

Specification	Linear	Binary	"Linear and district"
Aid presence, no lag	0.00529**	0.00489	0.00502**
	(0.00265)	(0.00640)	(0.00221)
Aid presence, one-year lag	0.00760**	0.00970	0.00478*
	(0.00341)	(0.00719)	(0.00289)
Aid presence, two-year lag	0.00639	0.00967	0.00366
	(0.00410)	(0.00835)	(0.00386)
Aid presence, three-year lag	0.00850	0.0105	0.00565
	(0.00776)	(0.0115)	(0.00639)
Aid presence, four-year lag	0.00529	0.00375	0.00415
	(0.0118)	(0.0150)	(0.00943)

Twenty-kilometer radius:

Specification	Linear	Binary	"Linear and district"
Aid presence, no lag	-0.00123	0.000108	-0.000374
	(0.00148)	(0.00501)	(0.00134)
Aid presence, one-year lag	-0.000920	0.00191	-0.00114
	(0.00178)	(0.00550)	(0.00163)
Aid presence, two-year lag	-0.00365*	-0.00443	-0.00399**
	(0.00209)	(0.00612)	(0.00202)
Aid presence, three-year lag	-0.00257	-0.000953	-0.00195
	(0.00426)	(0.00770)	(0.00367)
Aid presence, four-year lag	-0.000493	0.00607	-0.000158
	(0.00572)	(0.0102)	(0.00527)

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

5.3.4. Local Model Restricted to 2008

The birth records stem from questionnaires answered by the mothers. One could argue that the birth records, therefore, suffer from measurement error as the mothers do not remember exact dates of birth and death. One other issue is that mothers who have passed away are not recorded. To reduce both possible errors we run the model for observed births in 2008 only. It is the closest year to the survey date, and it should contain the least errors due to both reasons. The control variables are also more relevant in time as they are gathered at the time of the survey. The estimated coefficients of aid presence are presented in table 7. Note that when we only use data for one year, the specification for linear and "linear and district" is the same, as the district dummies captures the effect of the projects coded on district level.

Table 7 Local model restricted to year 2008

Radius	10-km		20-km	
Specification	Linear	Binary	Linear	Binary
Aid presence, no lag	0.00466	0.00780	-0.00396	-0.00556
	(0.00388)	(0.0127)	(0.00250)	(0.0101)
Aid presence, one-year lag	0.00634	0.00780	-0.00524*	-0.00482
	(0.00503)	(0.0127)	(0.00304)	(0.0100)
Aid presence, two-year lag	0.00592	0.00780	-0.00595*	-0.00482
	(0.00516)	(0.0127)	(0.00309)	(0.0100)
Aid presence, three-year lag	0.0104	0.0180	-0.00869	0.00274
	(0.0105)	(0.0177)	(0.00688)	(0.0118)
Aid presence, four-year lag	-0.00983	-0.0109	-0.0358**	-0.00681
	(0.0212)	(0.0283)	(0.0154)	(0.0217)

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

The result for year 2008 is a bit different compared to all years. For the ten-kilometer specification, there are no significant estimates and they are negative for the four-year lag. For the twenty-kilometer specification, all except one coefficient estimate are negative and the absolute magnitude of the estimates is larger. For the linear specification, the estimates are significant for the one-year lag, two-year lag and four-year lag. Overall, the result is not particular different for the ten-kilometer specification while the twenty-kilometer one implies stronger evidence for a negative effect of aid on infant mortality. The evidence is still very weak, as it only shows significant effects for the linear specification with a twenty-kilometer radius.

5.3.5. Instrumentation

One problem with the previous local model is that it does not address potential reverse-causality due to endogenous allocation of aid. One advantage with geocoded aid data is that it might provide means to overcome the problem of endogenous allocation. This geographical variable has to be decisive in the allocation of aid, but should not impact health outcomes directly. This variable would be an instrumental variable to estimate the effect of aid.

When examining the dataset, we find that most health aid projects with exact point locations are allocated to cities and especially district capitals. This rules out many instrumental variables based on geographical information. The degree of urbanization could, however, be a suitable instrument. This might also indicate that the collection of precise geocodes is systematically flawed. It is possible, that aid projects, which are not allocated to a specific district capital, are simply coded as projects on district level, as it is might be difficult to determine the exact location of projects in rural areas. Projects focusing on rural areas could also be active in many locations in the district and, therefore, coded at the district level without exact point locations.

In the DHS data the de facto residence of each cluster is recorded as urban or rural. In order to be a valid instrument, this variable has to fulfill the assumptions of exclusion and relevance. There is correlation of urbanization and health presence as shown in table 8 below. The relatively low values indicate that a substantial amount of observations is made in urbanized areas without presence of health aid projects. If added to our local model specifications, the urban dummy does not explain infant mortality significantly. It, therefore, seems reasonable, that the degree of urbanization itself does not have a direct impact on infant mortality. Though the exclusion assumption cannot be tested, this is support for the fulfillment of the exclusion assumption.

Table 8 Correlation of urban variable and health aid presence

	No lag	One-year lag	Two-year lag	Three-year lag	Four-year lag
Aid presence binary	0.2466	0.2140	0.1891	0.1431	0.1014
Aid presence linear	0.2863	0.2401	0.1985	0.1398	0.0817

For health aid presence within a twenty-kilometer radius the correlations are slightly lower, but similar, and therefore not presented. The results for instrumenting the health lags in both linear and binary form, both for ten-kilometer and twenty-kilometer, with the urban variable are presented in table 9 below. The coefficients of the controls do not show meaningful differences among these specifications, we do show the regression with controls for the one-year lag and a radius of ten kilometers in table B.V in appendix B and its first stage regression in table B.IV in appendix B.

Table 9 Instrumental variable model

	10-km radius		20-km radius	
Specification	Linear	Binary	Linear	Binary
No lag	0.00287	0.00693	0.00265	0.00823
	(0.0101)	(0.0243)	(0.00930)	(0.0288)
One-year lag	0.00430	0.00884	0.00409	0.0101
	(0.0151)	(0.0310)	(0.0143)	(0.0354)
Two-year lag	0.00665	0.0120	0.00643	0.0131
	(0.0233)	(0.0421)	(0.0225)	(0.0461)
Three-year lag	0.0170	0.0232	0.0146	0.0218
	(0.0596)	(0.0814)	(0.0511)	(0.0766)
Four-year lag	0.0442	0.0492	0.0309	0.0408
	(0.155)	(0.172)	(0.108)	(0.143)

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

As can be seen, the estimates are positive among all specifications, but this approach does not yield significant results. Apart from the interpretation that aid does not yield a decrease in infant mortality, it might be that urban is not explaining the allocation of aid projects well enough. However, the first stage results show strong significance for aid presence.

The urban variable is correlated with the health presence variables and shows significance in all first stage regressions, while the first stage regressions also all show a larger F-value than 15, which indicates relevance of the variable urban as an instrument. Otherwise most control variables also show significance in the first stage. Taken like this, the results suggest that infant mortality did not decrease in proximity of local health aid projects.

These results should, however, not be taken as strong evidence for health aid to be ineffective in decreasing infant mortality. It might be, that the effect on infant mortality is in reality small and negative, while the problems with systematically missing geocodes of precision one and the instrument urban not explaining aid well enough lead to the positive sign and insignificant results.

5.3.6. Matching

Matching is one other approach to solve the problem of non-random allocation of aid projects in the local model. The allocation of health aid projects can be seen as non-random allocation of a treatment. Matching can then be used to estimate causal effects, as long enough variables are observed that explain treatment. Besides health related questions, the DHS collects additional information about each respondent, such as education, ethnicity, wealth and nutritional conditions. Under the assumptions of conditional independence of potential outcomes and treatment on these covariates and common support, matching can be used to solve the problem of the arising selection bias (Caliendo, Kopeinig 2008). We discuss, whether these assumptions are reasonable in our setting and report results from nearest neighbor matching and propensity score matching.

In this setting, the treatment is proximity of a health aid project. The assumption of conditional independence is in principle a form of exogeneity (Angrist, Pischke 2008). In matching it states, that treatment status conditional on observable variables is independent from potential outcomes. In our case this means that potential infant mortality is independent from the presence of health aid projects, as long as we control for other variables in the DHS data. In other words, all differences in infant mortality between the observed births close and not close to health aid projects, which are not depending on the health aid projects, can be controlled for with the available covariates. The criteria to pick variables for matching are that "only variables that simultaneously influence the treatment and the outcome should be included" (Caliendo, Kopeinig 2008). Furthermore, only variables not impacted by the treatment should be included. This is problematic in our setting, as the data is collected at the end of the considered time span. Therefore, we can include non-health related variables and fixed variables, which were not influenced by the health projects. This, however, might be problematic in case health aid projects are indeed allocated to areas most in need of health care. Too many non-relevant covariates increase the variance of the estimators and one should stick to parsimony when choosing the covariates for matching (Caliendo, Kopeinig 2008). However, it is still recommended to use as many covariates as reasonable and include them when in doubt, with regards to the assumption of conditional independence. Variables we included were district dummies, closeness to urban settlements, malaria endemicity and variables on individual level for each birth. For example maternal education was included, because it is a good predictor of infant mortality, socio-economic status and the proximity to health projects, as education is more widespread in the cities, where the projects are allocated. Ethnicities are also included, as the ethnical composition shows geographical variation and health delivery quality and institutions might vary among areas with different ethnical compositions. A problematic variable is HIV-status, which we excluded, because the proximity to health projects might impact it. This would pose a problem, if allocation of aid projects is highly correlated with HIV, because of a violation of the assumptions of conditional independence, which is not the case in our data. Other covariates were picked as they showed significant explanatory power in explaining health presence and it seemed plausible that they show differences among regions and have an impact on infant mortality. Table B.VI in appendix B presents the used variables. In cases in which several covariates are regarded, exact matching leads to only very few matches, because the number of possible combinations increases dramatically with each new control variable. However, using only one or fewer covariates for matching would violate the conditional independence assumption, as the treatment still depend on the omitted variables. For this reason Mahalanobis matching and propensity score matching are used. Table 10 presents the results for nearest neighbor matching, using the Mahalanobis metric as distance measure.

These results show significant negative effects of closeness to health projects for no lag, and a lag of four-year for aid projects within a ten-kilometer range and for the first three lags for a twenty-kilometer range. However, closest neighbor matching does not allow testing the assumption of common support as propensity score matching does. Furthermore, the Mahalanobis metric is used to match closest neighbors conditional on multiple covariates. Problematic with this metric is, that it puts the same weight on all variables (Stuart 2010). Beyond this, it has been shown that closest neighbor matching based on this metric, can yield biased results with more than eight variables and with non-normally distributed variables (Stuart 2010). Rosenbaum and Rubin (1983) showed that if the conditional independence assumption holds for the included covariates, it should also hold conditional on the propensity score constructed with these covariates. Propensity score matching should, therefore, yield similar results as nearest neighbor matching, if the conditional independence assumption holds. Furthermore, the explicit modeling of treatment probability allows analyzing whether the common support condition holds. We therefore also apply propensity score matching to obtain a better understanding of the validity of the above results.

Table 10 Nearest neighbor matching with Mahalanobis distance metric

Aid presence	10-km	20-km
No lag	-0.0135*	-0.0190***
	(0.00769)	(0.00491)
One-year lag	-0.0128	-0.0223***
	(0.00825)	(0.00595)
Two-year lag	-0.0109	-0.0264***
	(0.0107)	(0.00677)
Three-year lag	-0.00145	-0.0118
	(0.0152)	(0.00909)
Four-year lag	-0.0351***	0.00330
	(0.0104)	(0.0141)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

In propensity score matching the probability of treatment for all treated and non-treated observations is estimated. Treated and non-treated observations are then matched for likewise probability of treatment in order to estimate the average treatment effect or the average treatment effect on the treated. While there are several algorithms to select which observations exactly are compared, we stick with nearest neighbor matching based on the propensity score in this step, other algorithms are frequently used to check for robustness (Stuart 2010). Our treatment is a discrete variable and in order to estimate the propensity score properly, a model for propensity score estimation has to be chosen and the explanatory covariates have to be selected. Logit or Probit are commonly used for propensity score estimation, as they do not predict negative probabilities or probabilities larger than one. We choose a Logit estimate but the effect of this choice on the result is negligible. The covariates included to estimate the propensity are the same as described for nearest neighbor matching. When in doubt whether to include them, parsimony was followed in fitting the Logit estimate, as suggested by Caliendo and Kopeinig (2008).

Table 11 Propensity score matching

Aid presence	10-km	20-km
No lag	8.63e-05	0.0258*
	(0.132)	(0.0151)
One-year lag	-0.0217	0.00337
	(0.0139)	(0.0185)
Two-year lag	0.177	-0.0175
	(0.122)	(0.0116)
Three-year lag	-0.0238	-0.0211***
	(0.0199)	(0.00801)
Four-year lag	-0.0450**	-0.0129
	(0.0179)	(0.0198)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The results of propensity score matching are shown in table 11. For the three different lags the specifications have to be changed to some extent. This is the case, because for higher lags, fewer observations show health presence. Then, subgroups of our data become perfect predictors of no health presence, which means they cannot be used for a Logit or Probit estimate of treatment probability. Hence, these observations, or these variables, are dropped, depending on how many observations are affected – if the variable would lose its meaningfulness, it is dropped. If only few observations are affected, they are dropped and the variable is kept. Problematic are the specifications for the lags three and four years of aid presence. Here, most of the districts would have to be dropped, and hence the three larger regions are used instead of district dummies. It is therefore questionable, whether the conditional independence assumption can be assumed to hold under these conditions, because there is substantial variation in infant mortality within the regions. For completeness, the specification with regions was also used for no and one lag, but the results do not change meaningfully. Furthermore, some observations would violate the overlap assumption in this specification of the propensity score. We follow Caliendo and Kopeinig (2008) and delete the non-overlapping observations, when applying the propensity score estimation.

Common support or overlap requires that both treated and untreated observations have a probability larger than zero and smaller than one of being treated. This ensures that a counterfactual observation for each treated individual can be found. Furthermore, for a meaningful estimation of the average treatment effect, enough probability mass has to overlap, so that comparisons between truly similar observations can be made.

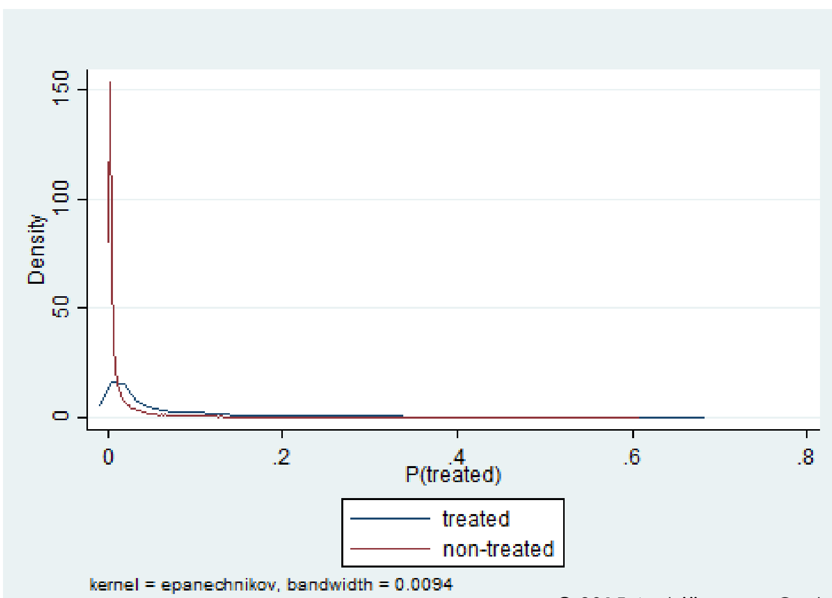
Whether the overlap assumption holds can be discussed by examining the used propensity score. We present a plot, figure 1, of estimated density for treated and non-treated for the significant result, with the fourth lag of health aid presence as a treatment. Evidently, the largest part of the probability mass is allocated around zero. This is problematic, because it implies that the propensity score does not fulfill the assumption of common support for many observations. Though there is still overlap for higher probabilities than under one percent, this is the case for only 5,259 of the 18,032 observations used. Hence, there is serious doubt that common support is sufficient in our data.

The main goal of the propensity score is to balance the covariates, but the predictive power of the propensity score gives another clue about the quality of its specification (Heinrich, Maffioli & Vazquez 2010). The predictive power decreases with increasing lags of health presence. For the fourth lag of health aid presence only 16.62 % of the observations have a propensity score higher than 50 %. Though the conditional independence assumption cannot be tested and it still could be fulfilled, this suggests that the propensity score suffers from omitted variable bias. This would imply a violation of the conditional independence assumption. Taken together this suggests that the result for the fourth lag is not a robust result and that propensity score matching in this form does not yield meaningful results.

Propensity score matching and nearest neighbor matching with Mahalanobis metric rely on the same assumptions and should yield at least similar results (Rosenbaum, Rubin 1983). Both methods are conditioning on the same covariates and considering the problems of nearest neighbor matching mentioned above, we have to conclude that the first results obtained by nearest neighbor matching are not valid. However, we believe that richer data and a better understanding of aid allocation would make propensity score matching a viable approach to measure the effect of aid presence on infant mortality with this method.

The propensity score estimated here is also an attempt to model aid allocation. This suggests that the non-health covariates in the DHS data are not sufficient to estimate the allocation of health aid projects, i.e. that there are omitted variables. However one can conclude that the allocation of aid is non-random, because several variables show significant effects when estimating the propensity score. We present the estimate of the propensity score for no lag and aid presence within ten-kilometer range in table B.VII in appendix B.

Figure 1 Kernel density plot of propensity score



6. Discussion and Conclusion

On the basis of our results we cannot reject the null hypothesis that aid projects directed at the health sector do not decrease infant mortality in Malawi. However, several caveats and difficulties with our approaches are discussed earlier, and this result should be taken with caution. In this section we discuss the validity of our results and point to possible improvements.

For the two district models, we find small positive effects when aid presence is specified as the number of projects in the districts. When using commitment or disbursement per capita as aid presence instead, the effects are negative for few lags and positive for more lags. Overall few estimates are significant. This together with the different signs for the specifications makes the result inconclusive.

The model on local level looks at health aid projects in close proximity to observed births. This might be a way to detect small-area patterns below district level, which might be lost at aggregate level. Due to the random displacement of the DHS data we cannot choose very small areas and we look at radiuses of ten and twenty kilometers around the births. The ten-kilometer specification shows small positive effects, whereas most of the estimated coefficients for twenty-kilometer is negative and have smaller absolute magnitudes. The different results for the two chosen radiuses are unexpected. One potential explanation for the larger number of negative estimates for the twenty-kilometer radius specification, could be that the ten kilometer-radius is too small to cover all individuals that are actually affected by the project, this would bias the estimates for the ten-kilometer specification. On the other hand, the smaller absolute magnitude for the twenty-kilometer specification might be, because the area includes projects too far away from the mother to have an effect on infant mortality. This would also bias the estimates. In order to investigate this further it is necessary to obtain more detailed information about the projects and their reach. Most estimates for both radiuses are however insignificant. One complication with analyzing the data on local level is small sample size. One approach to deal with this problem could be to use more advanced methods that take into account the geo-spatial correlation between small areas. The mortality would then be estimated by borrowing information across neighboring areas, see for example Chin et al. (2011).

In order to try and account for potential reverse causality in the local model an instrumental variable approach is applied. We used urbanization as an instrument and find no significant effects, while the indicator of urbanization shows significance in the first stage regression. It is possible, that a better instrument can be found, which should explain the allocation of aid better and lead to more reliable results. However, the allocation of health aid projects mostly to district capitals makes it difficult to find a geographical instrument. In conclusion, while this result seems to add to the evidence of no impact of health aid projects on infant mortality, it is possible, that the instrument does not explain aid allocation well enough and that a better instrument finds significant results.

Matching is also utilized in the local model as it seems to be a viable approach, with the rich data available in the DHS, to solve the empirical problem of non-random allocation of treatment. At a closer look, however, the data does not seem to fulfill the necessary assumptions of conditional independence of treatment and common support. Therefore, the significant results of Mahalanobis matching and the non-significant results of propensity score matching should both not be taken at face value.

The internal validity of the result that aid projects do not at all decrease infant mortality in Malawi has to be doubted for the reasons stated above. However, the effects of most local aid projects cannot be very large, as we otherwise should have found more indications of an effect on infant mortality. That is unless the aid presence data is seriously flawed and many efficient projects are not geocoded. In terms of external validity this study does not provide conclusions for other countries, other than aid being allocated non-randomly on a sub-national level, which should be considered for any other attempt to assess aid effectiveness.

Even if we cannot find a negative effect of aid presence on infant mortality, it is declining worldwide and also rapidly in Malawi. One can see this in our results when looking at the declining trend of the yearly dummies and also in World Bank data, according to which it declined from 90 to 59 between 2002 and 2008 (The World Bank, UN Inter-agency Group for Child Mortality Estimation). The exact drivers of this development seem to be of a nationwide nature, as our results suggest. The effectiveness and impact of these drivers are questions for future research. Liu et al. (2012) and Mustafa and Odimegwu (2008) suggest that important drivers are bed-nets to prevent malaria infections and other interventions against infectious diseases. We also used DHS data as Mustafa and Odimegwu (2008) but do not find a significant effect of the variable "slept last night under a bednet" in various specifications, such as an interaction with malaria endemicity. If these interventions are not the most important driver one other plausible factor is general improvement of the public health care system in Malawi. Our data does not allow at this point to test this. In assessing the overall effectiveness of aid in the health care sector, it would be important to test how effectively aid assisted in improving health delivery on a nation-wide level.

The results of the instrumental variable regression and propensity score matching are interesting in a different respect. Both the first stage regression and the propensity score show that aid projects are not allocated randomly, but also not clearly because of need. This is also evident when examining the data. Almost all health aid projects are allocated to the district capitals. This makes several variables good candidates for the explanation of allocation. Geographic accessibility and administrative convenience, which should be the highest in the most developed locations within a country, are most likely drivers. One other explanation is that in the urban centers, simply more people can be reached, which makes the aid projects potentially more efficient. Furthermore, different sectors seem to have very different allocation rules, which might be a field for interesting research.

The analyses summarized above are very dependent on the geocoded aid data. The data for Malawi used is the most extensive available at present for any country. The geocoded datasets are however continuously updated and data for more countries is made available by time. With better and more reliable aid data there is potential for improved studies in this area in the future. Our analysis is also limited by the fact that the latest DHS for Malawi was conducted in 2010, forcing us to end the study in 2008. Access to a later DHS or other birth records, would increase the sample size, potentially increasing the power of the study.

Even with access to more comprehensive geocoded datasets, the data pose some limitations and uncertainties. Ideally one would like to know exact location, coverage and the time span a project is active, not only when the agreement is signed. It would also be good to know the disbursement and commitment per location. This lack of information is most likely the result of the goal by AidData to code a comprehensive number of projects per country instead of focusing on fewer projects with more detailed descriptions. In order to further investigate the effect of local aid projects on health outcomes it could be necessary to collect more detailed information about the projects. This is probably necessary to do locally. More detailed data would also help to select which health projects to use. More information about the projects would allow selecting projects that target infant mortality directly. One could then differentiate the time it takes for different projects to affect mortality. In this context the precision of the coding should also be mentioned, because most health aid projects with exact point locations are located in cities. Some other aid sectors in the dataset have, however, more project locations outside the cities. This indicates that the geocoding method can code exact locations in rural areas properly. But one should still remain skeptical of the concentration of health projects with exact locations in cities.

Summing up, local aid projects directed at health do not seem to have a significantly decreasing effect on infant mortality, according to our findings. However, there remain problems with data and endogeneity, which pose questions for future research. Furthermore, aid projects with a local impact are not all of the aid that is directed at the health sector in Malawi and certain nation-wide interventions are not covered by this approach. These should be taken into account before assessing the overall effectiveness of aid in the health sector in Malawi.

7. Bibliography

Alesina, A. & Dollar, D. 2000, "Who gives foreign aid to whom and why?" *Journal of Economic Growth,* vol. 5, no. 1, pp. 33-63.

Angrist, J.D. & Pischke, J. 2008, *Mostly Harmless Econometrics: An Empiricist's Companion,* Princeton University Press, Princenton, NJ

Baird, S. Hicks, J.H., Kremer, M. & Miguel, E. 2011, "Worms at work: long-run impacts of child health gains," unpublished paper, [October 2011].

Banister, J. & Preston, S.H. 1981, "Mortality in China," *Population and Development Review,* vol. 7, no. 1, pp. 98-110.

Bazzi, S. & Clemens, M.A. 2013, "Blunt instruments: Avoiding common pitfalls in identifying the causes of economic growth," *American Economic Journal: Macroeconomics,* vol. 5, no. 2, pp. 152-186.

Boone, P. 1996, "Politics and the effectiveness of foreign aid," *European Economic Review,* vol. 40, no. 2, pp. 289-329.

Bourguignon, F. & Sundberg, M. 2007, "Aid effectiveness: opening the black box," *American Economic Review,* vol. 97, no. 2, pp. 316-321.

Burnside, C. & Dollar, D. 2000, "Aid, growth and policies," *American Economic Review,* vol. 90, no. 4, pp. 847-868.

Caliendo, M. & Kopeinig, S. 2008, "Some practical guidance for the implementation of propensity score matching," *Journal of Economic Surveys,* vol. 22, no. 1, pp. 31-72.

Chin, B., Montana, L. & Basagaña, X. 2011, "Spatial modeling of geographic inequalities in infant and child mortality across Nepal," *Health & Place,* vol. 17, no. 4, pp. 929-936.

Collier, P. 2006, "Is aid oil? An analysis of whether Africa can absorb more aid," *World Development,* vol. 34, no. 9, pp. 1482-1497.

Cutler, D.M., Deaton, A.S. & Lleras-Muney, A. 2006, "The Determinants of mortality," *Journal of Economic Perspectives*, vol. 20, no. 3, pp. 97-120.

De, R. & Becker, C. 2015, "The foreign aid effectiveness debate: Evidence from Malawi," *AidData Working Paper Series*, no. 6, AidData, Washington DC.

Demombynes, G. & Trommlerová, S.K. 2012, "What has driven the decline of infant mortality in Kenya?" *World Bank Policy Research Working Paper,* no. 6057, World Bank, Washington DC.

Easterly, W. 2007, "Was development assistance a mistake?" *American Economic Review,* vol. 97, no. 2, pp. 328-332.

Fogel, R.W. 1997, "New findings on secular trends in nutrition and mortality: some implications for population theory," *Handbook of Population and Family Economics*, eds. M.R. Rosenzweig & O. Stark, Elsevier Science Publishers, Amsterdam.

Frot, E. & Perrotta, M. 2012, "Aid effectiveness: new instrument, new results," *SITE Working Paper Series*, vol. 11, Stockholm Institute of Transition Economics, Stockholm.

Gething, P.W., Patil, A.P., Smith, D.L., Guerra, C.A., Elyazar, I., Johnston, G.L., Tatem, A.J. & Hay, S.I. 2011, "A new world malaria map: Plasmodium falciparum endemicity in 2010," *Malaria Journal,* vol. 10, no. 378, pp. 1475-2875.

Greenstone, M. & Hanna, R. 2011, "Environmental regulations, air and water pollution, and infant mortality in India," *NBER Working Papers*, no. 17210, National Bureau of Economic Research, Cambridge, MA.

Heinrich, C., Maffioli, A. & Vazquez, G. 2010, *A primer for applying propensity-score matching*, Office of Strategic Planning and Development Effectiveness, Impact Evaluation Guidelines, Technical Notes, no. IDB-TN-161, Inter-American Development Bank, Washington DC.

Liu, L., Johnson, H.L., Cousens, S., Perin, J., Scott, S., Lawn, J.E., Rudan, I., Campbell, H., Cibulskis, R. & Li, M. 2012, "Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000," *The Lancet,* vol. 379, no. 9832, pp. 2151-2161.

Lumsdaine, D.H. 1993, *Moral Vision in International Politics: The Foreign Aid Regime, 1949-1989,* Princeton University Press, Princeton, NJ.

Malawi National Statistical Office 2008*, 2008 Population and Housing Census - Preliminary report*. Available: http://www.malawihighcommission.co.uk/MWCensus08.pdf [2015, 05/06].

Miguel, E. & Kremer, M. 2004, "Worms: identifying impacts on education and health in the presence of treatment externalities," *Econometrica,* vol. 72, no. 1, pp. 159-217.

Mishra, P. & Newhouse, D.L. 2007, "Health aid and infant mortality," *IMF Working Paper*, no. 07/100, International Moneteray Fund, Washington DC.

Mustafa, H.E. & Odimegwu, C. 2008, "Socioeconomic determinants of infant mortality in Kenya: analysis of Kenya DHS 2003," *Journal of Humanities and Social Science,* vol. 2, no. 2, pp. 1934-1722.

National Statistical Office (NSO) and ICF Macro 2015, *Malawi Demographic and Health Survey 2010,* NSO and ICF Macro, Zomba, Malawi, and Calverton, Maryland, USA.

Peratsakis, C., Powell, J., Findley, M., Baker, J. & Weaver, C. 2012, *Geocoded Activity-Level Data from the Government of Malawi's Aid Management Platform*, AidData and the Robert S. Strauss Center for International Security and Law*,* Washington DC.

Pritchett, L. & Summers, L.H. 1996, "Wealthier is healthier," *Journal of Human Resources*, vol. 31, no. 4, pp. 841-868.

Rajan, R.G. & Subramanian, A. 2008, "Aid and growth: What does the cross-country evidence really show?" *Review of Economics and Statistics,* vol. 90, no. 4, pp. 643-665.

Rogerson, A. & Steensen, S. 2009, "Aid orphans: whose responsibility," *Development Brief,* vol. 1, Organisation for Economic Co-operation and Development, Paris.

Rosenbaum, P.R. & Rubin, D.B. 1983, "The central role of the propensity score in observational studies for causal effects," *Biometrika,* vol. 70, no. 1, pp. 41-55.

Schultz, T.P. 2004, "School subsidies for the poor: Evaluating a Mexican Progresa poverty program," *Journal of Development Economics*, vol. 74, no. 1, pp. 199-250.

Spears, D. 2012, "Policy lessons from implementing India's total sanitation campaign," *India Policy Forum 2012*, vol.9, no. 1.

Strandow, D., Findley, M., Nielson, D. & Powell, J. 2011, "The UCDP and AidData codebook on georeferencing aid: Version 1.1," *Uppsala Conflict Data Program Paper Series,* no. 4, Department of Peace and Conflict Research, Uppsala University, Uppsala.

Stuart, E.A. 2010, "Matching methods for causal inference: A review and a look forward," *Statistical Science : A Review Journal of the Institute of Mathematical Statistics,* vol. 25, no. 1, pp. 1-21.

Svensson, J. & Björkman, M. 2009, "Power to the people: Evidence from a randomized field experiment of a community-based monitoring project in Uganda," *Quarterly Journal of Economics,* vol. 124, no. 2, pp. 735-769.

World Bank, *World Development Indicators*. Available: http://data.worldbank.org/country/malawi [2015, 5/5].

World Bank, UN Inter-agency Group for Child Mortality Estimation*, World Development Indicators - Infant Mortality Rate*. Available: http://data.worldbank.org/indicator/SP.DYN.IMRT.IN/countries/1W-MW?page=2&display=default [2015, 5/9].

Wooldridge, J. 2012, *Introductory Econometrics: A Modern Approach,* Cengage Learning, Andover.

Yontcheva, B. & Masud, N. 2005, "Does foreign aid reduce poverty? Empirical evidence from nongovernmental and bilateral aid," *IMF Working Papers*, no. 05/100, International Monetary Fund, Washington DC.

Appendix A: Data Description

Table A.I Summary statistics of the yearly cohorts in the DHS.

Year	Observed births	Observed infant deaths	Infant mortality rate
2002	2,371	244	102.9
2003	3,156	254	80.5
2004	3,352	262	78.2
2005	2,950	180	61.0
2006	3,063	221	72.2
2007	3,349	190	56.7
2008	3,543	230	64.9
Total	21,784	1,581	72.6

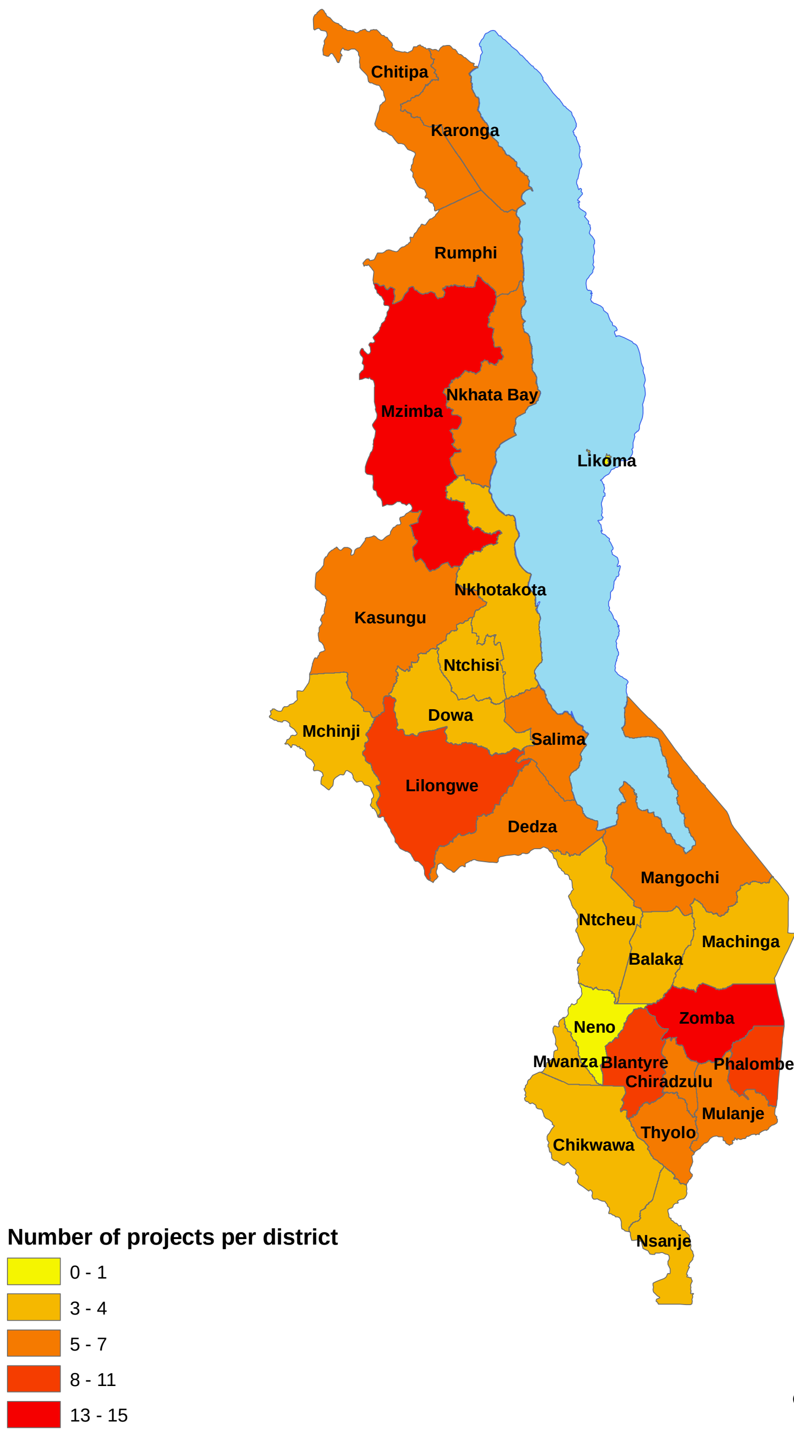
Table A.II Summary statistics of the districts in Malawi.

District	Region	Disburse-ment per capita ($)	Commit-ment per capita ($)	Projects with exact point locations	Projects at district level or lower	Population in 2008	Observed infant deaths	Observed births	Infant mortality
Balaka	Southern	6.3	7.8	2	4	316,748	61	872	70.0
Blantyre	Southern	7.2	9.4	6	9	999,491	45	607	74.1
Chikwawa	Southern	4.0	4.2	0	4	438,895	54	770	70.1
Chiradzulu	Southern	67.6	57.1	3	5	290,946	55	790	69.6
Chitipa	Northern	27.7	23.1	3	6	179,072	38	712	53.4
Dedza	Central	2.8	2.3	3	5	623,789	59	864	68.3
Dowa	Central	1.4	1.5	1	3	556,678	38	713	53.3
Karonga	Northern	7.6	5.5	4	6	272,789	24	637	37.7
Kasungu	Central	5.6	6.1	2	5	616,085	63	799	78.8
Lilongwe	Central	4.0	3.3	7	11	1,897,167	56	841	66.6
Machinga	Southern	2.4	3.0	2	4	488,996	70	920	76.1
Mangochi	Southern	7.2	10.7	3	7	803,602	76	905	84.0
Mchinji	Central	1.8	1.8	1	3	456,558	55	858	64.1
Mulanje	Southern	5.5	4.4	3	6	525,429	74	889	83.2
Mwanza	Southern	11.8	9.3	2	4	94,476	56	759	73.8
Mzimba	Northern	9.4	8.7	9	13	853,305	64	809	79.1
Neno	Southern	1.4	0.0	0	1	108,897	62	801	77.4
Nkhatabay	Northern	10.1	7.5	4	7	224,224	52	583	89.2
Nkhota Kota	Central	3.5	4.8	2	4	301,868	43	815	52.8
Nsanje	Southern	12.8	13.2	2	4	238,089	72	832	86.5
Ntcheu	Central	2.3	1.8	2	4	474,464	77	963	80.0
Ntchisi	Central	4.7	6.5	2	3	224,098	51	831	61.4
Phalombe	Southern	35.4	48.5	5	8	313,227	98	1058	92.6
Rumphi	Northern	10.9	5.2	3	6	169,112	33	657	50.2
Salima	Central	5.9	7.2	2	5	340,327	69	799	86.4
Thyolo	Southern	5.1	5.7	3	6	587,455	69	859	80.3
Zomba	Southern	17.5	16.0	8	15	670,533	67	841	79.7
Total		8.5	8.5	84	158	13,066,320	1581	21784	72.6

Table A.III Summary statistics of categorical variables.

Variable	**Total number of observations**	**Share (%)**
Maternal education	21,784	
No education		21.16
Primary		69.38
Secondary and higher		9.46
Maternal age at birth	21,784	
Less than 20 years		18.69
20-less than 35 years		68.67
35 years or more		12.64
Sex of child	21,784	
Male		49.55
Female		50.45
Birth order and interval between births	21,142	
First child		18.05
2-4 / 2 years or more		40.83
5+ / 2 years or more		8.76
2-4/ less than 2 years		26.17
5+ / less than two years		6.19
Multiple birth	21,784	
No multiple birth		95.82
Multiple birth		4.18
HIV	18,033	
HIV-negative mother		94.12
HIV-positive mother		5.88
Urban	21,784	
Resides in urban area		6.48
Resides in rural area		93.52

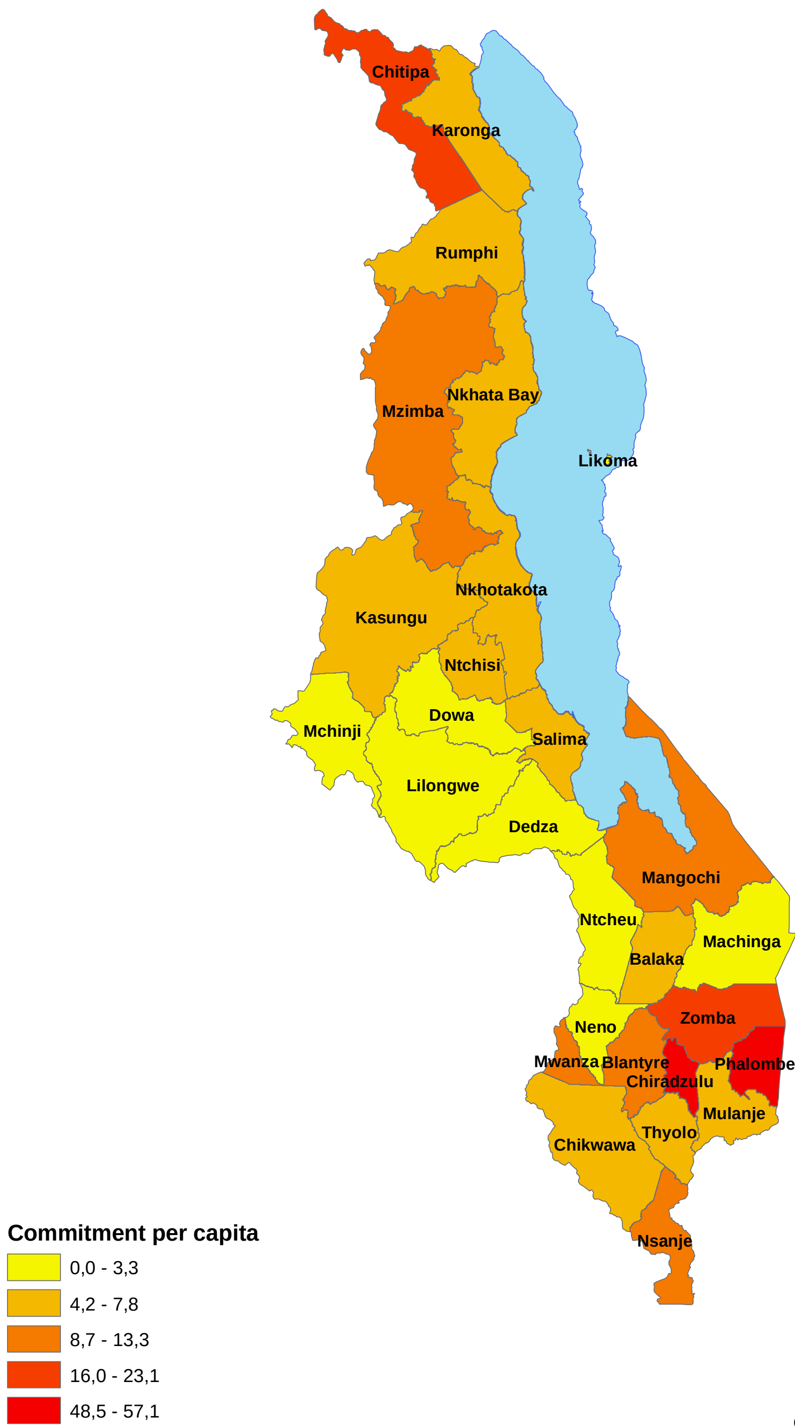
Figure A.I Health aid projects per district, year 2002-2008.



Based on own calculations.

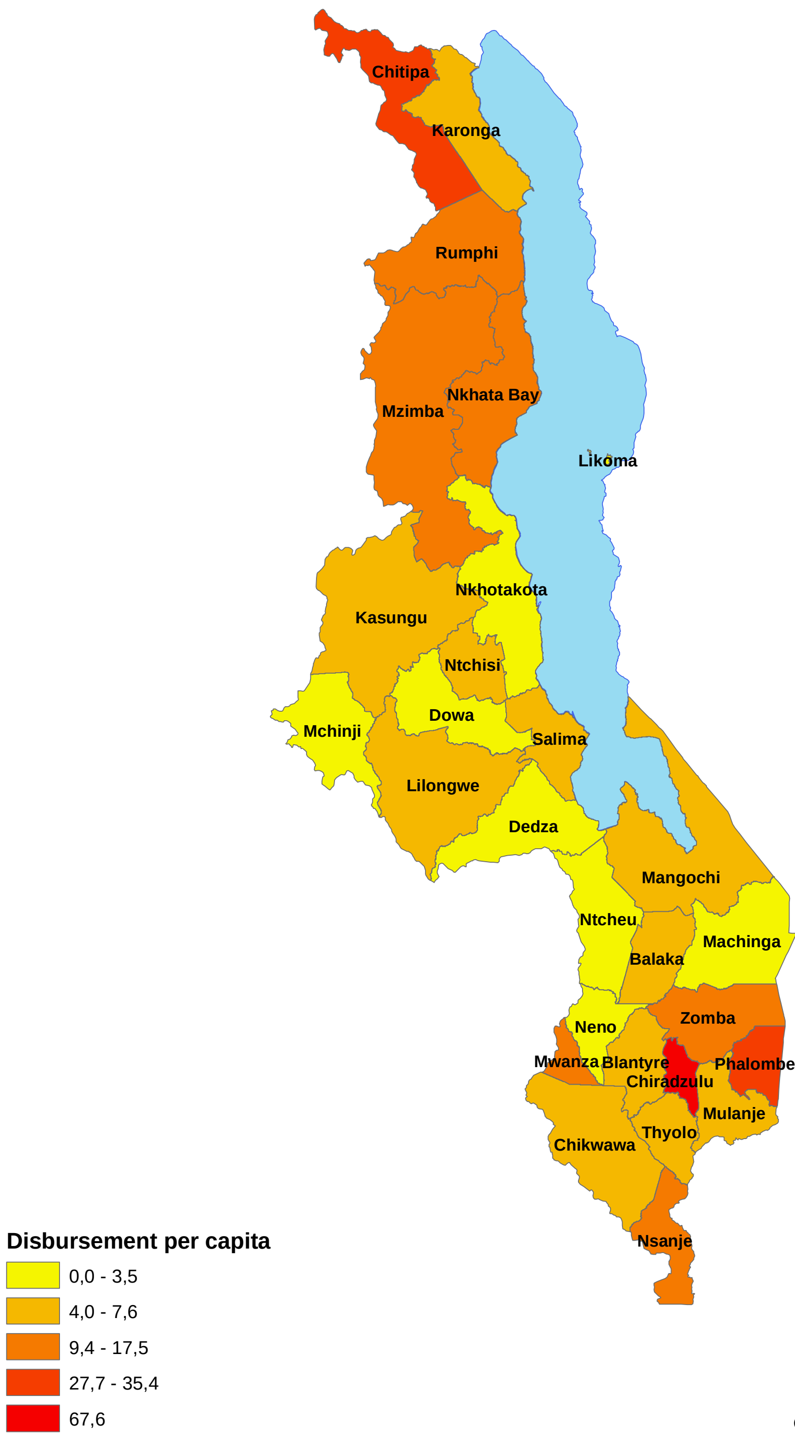
Data from Petraskis et al. (2012).

Figure A.II Commitment per capita of health aid projects per district, year 2002-2008.



Based on own calculations.
Data from Petraskis et al. (2012).

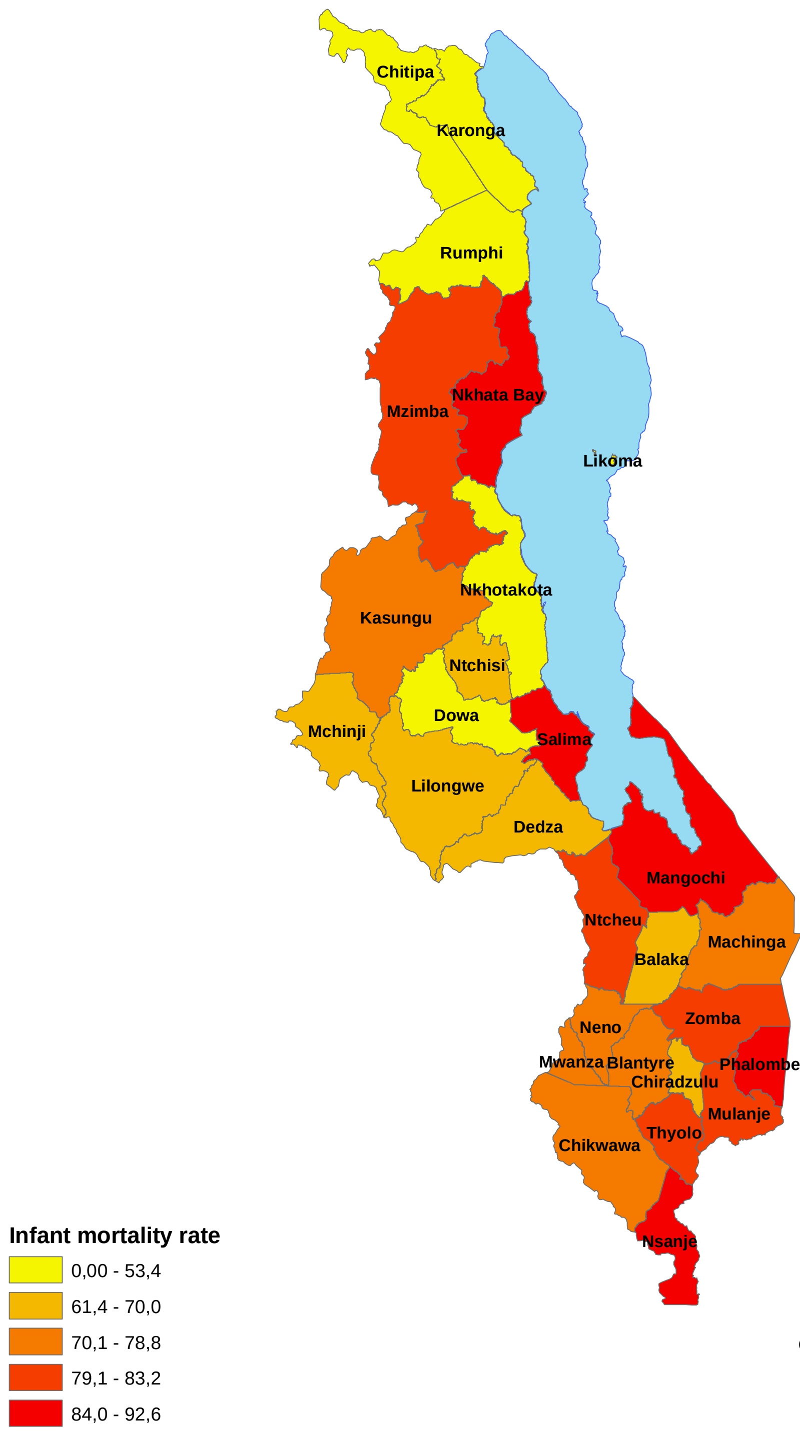
Figure A.III Disbursement per capita of health aid projects per district, year 2002-2008.



Based on own calculations.

Data from Petraskis et al. (2012).

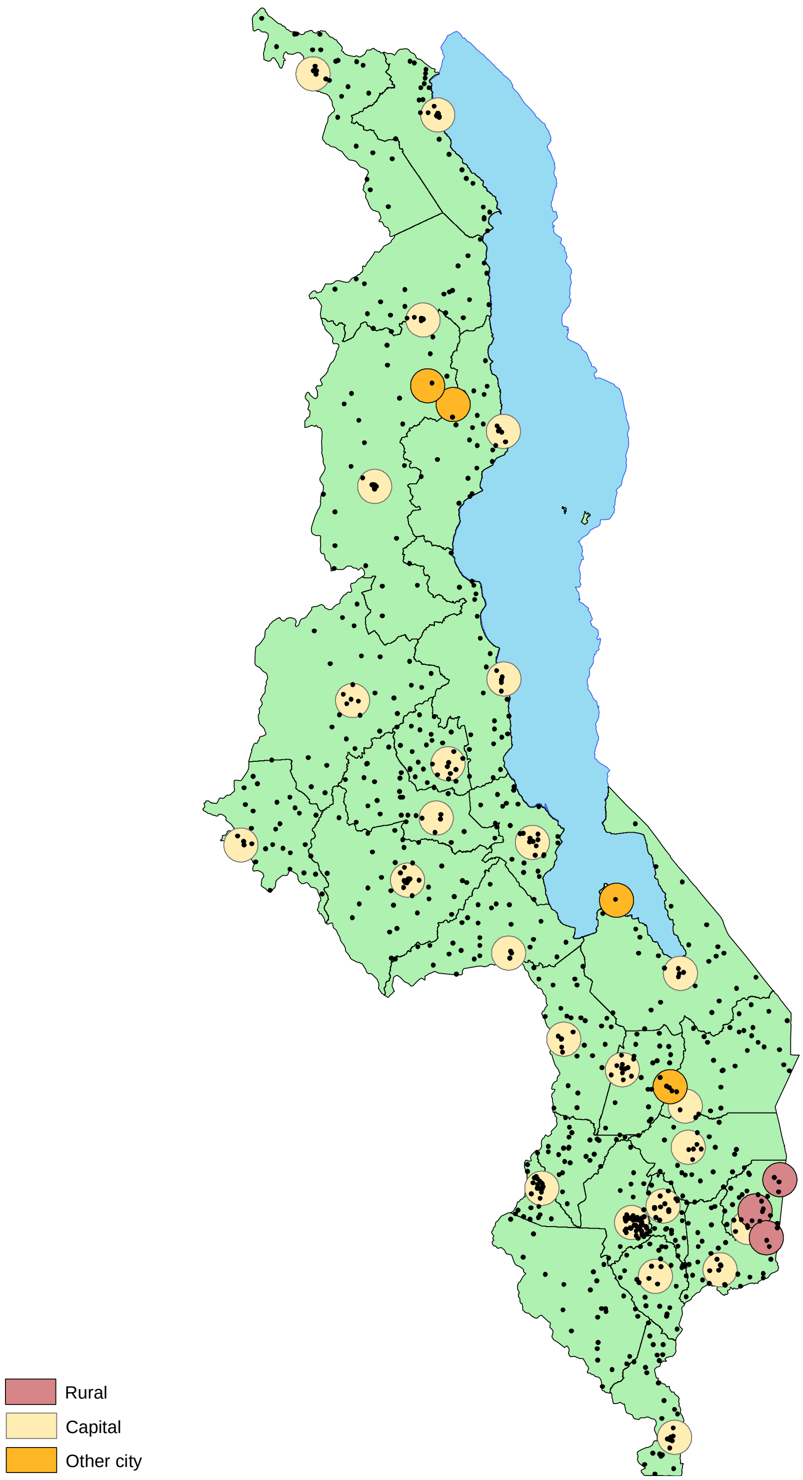
Figure A.IV Infant mortality rate per district, year 2002-2008.



Based on own calculations.

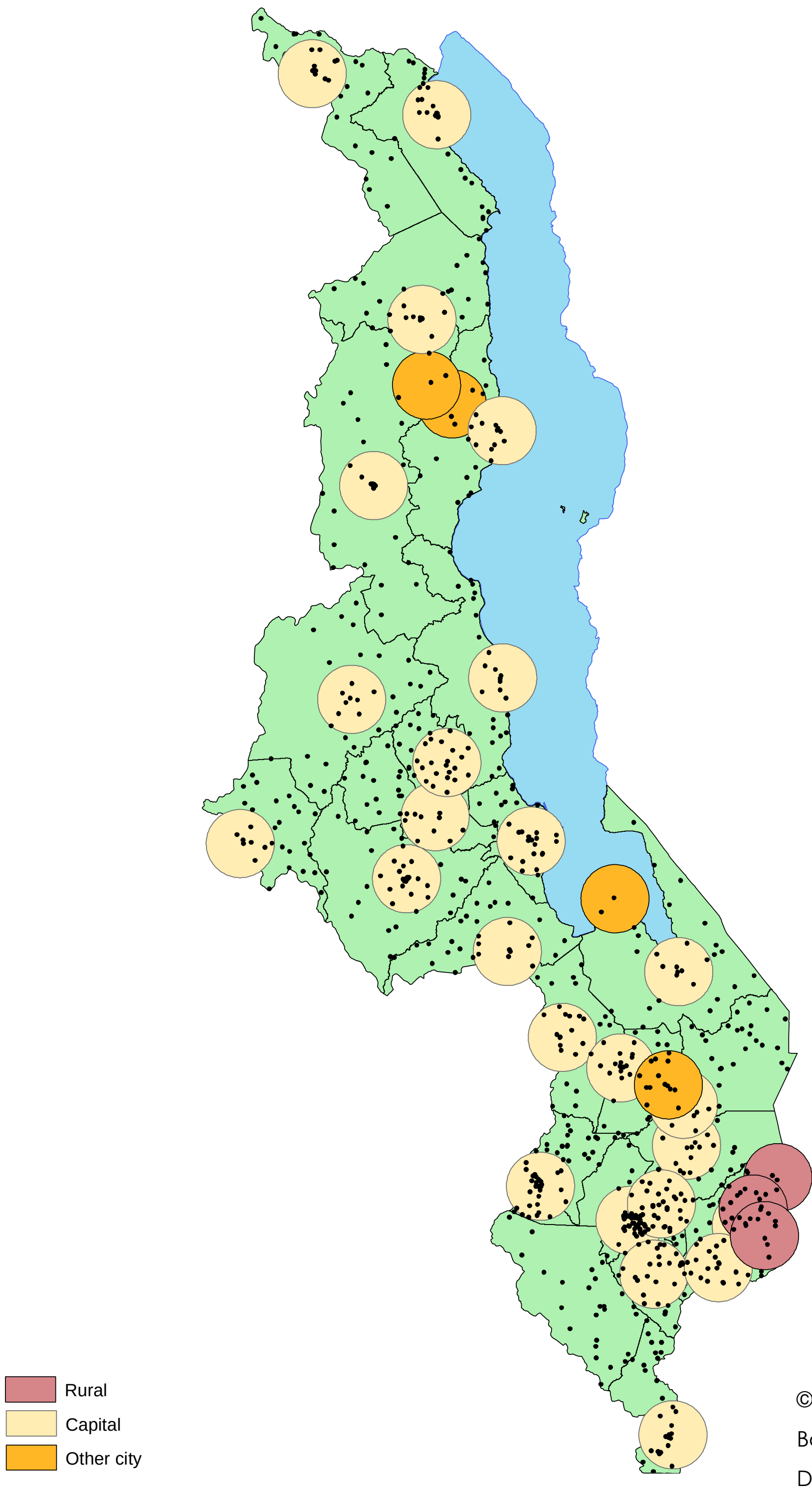
Data from National Statistics Office and ICF Macro (2015).

Figure A.V Aid project locations and their corresponding ten kilometers radiuses. Locations in rural areas are red, district capitals yellow and other cities orange. The black dots symbolize clusters in the DHS.



Based on own calculations.
Data from National Statistics Office and ICF Macro (2015), Petraskis et al. (2012).

Figure A.VI Aid project locations and their corresponding twenty kilometers radiuses. Locations in rural areas are red, district capitals yellow and other cities orange. The black dots symbolize clusters in the DHS.



Based on own calculations.
Data from National Statistics Office, ICF Macro (2015) and Petraskis et al. (2012).

Appendix B: Regression Tables

In the regression tables below district fixed effects are jointly significant, they are omitted due to limited space.

Table B.I Aggregated district model, aid presence specification: number of projects and no lag.

DEPENDENT VARIABLE	Infant mortality
Aid presence no, lag	4.276**
	(1.643)
Year (2002 base category)	
2003	-24.52***
	(6.950)
2004	-25.26***
	(7.308)
2005	-45.93***
	(6.687)
2006	-38.53***
	(8.169)
2007	-62.95***
	(9.891)
2008	-61.36***
	(11.22)
Constant	81.51***
	(7.010)
Observations	189
R-squared	0.421

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

Table B.II Individual district model, aid presence specification: number of projects and no lag.

DEPENDENT VARIABLE	Infant death
Aid presence, no lag	0.00285
	(0.00196)
Maternal education (no education base category)	
Primary	0.0111**
	(0.00498)
Secondary and higher	-0.00194
	(0.00779)
Maternal age at birth (less than 20 years base category)	
20-less than 35 years	-0.0168*
	(0.00863)
35 years or more	-0.00543
	(0.0106)
Sex of child (male base category)	
Female	-0.0122***
	(0.00377)
Birth order and interval between births (first order base category)	
Order 2-4 / 2 years or more interval	-0.0214**
	(0.00847)
Order 2-4 / less than 2 years interval	0.0444***
	(0.0124)
Order 5+ / 2 years or more interval	-0.0243**
	(0.0104)
Order 5+ / less than 2 years interval	0.0627***
	(0.0126)
Multiple birth	0.187***
	(0.0204)
HIV-positive mother	0.0930***
	(0.0101)
Year (2002 base category)	
2003	-0.0248***
	(0.00847)
2004	-0.0333***
	(0.00803)
2005	-0.0426***
	(0.00816)
2006	-0.0365***
	(0.0115)
2007	-0.0518***
	(0.0141)
2008	-0.0522***
	(0.0133)
Constant	0.101***
	(0.0101)
Observations	17,484
R-squared	0.043

Standard errors, clustered on district level, in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table B.III Local model, year 2002-2008, aid presence specification: linear, 10-km radius and no lag.

DEPENDENT VARIABLE	Infant death
Aid presence, no lag	0.00529**
	(0.00265)
Maternal education (no education base category)	
Primary	0.0107**
	(0.00530)
Secondary and higher	-0.00339
	(0.00781)
Maternal age at birth (less than 20 years base category)	
20-less than 35 years	-0.0168**
	(0.00686)
35 years or more	-0.00564
	(0.00977)
Sex of child (male base category)	
Female	-0.0120***
	(0.00387)
Birth order and interval between births (first order base category)	
Order 2-4 / 2 years or more interval	-0.0215***
	(0.00680)
Order 2-4 / less than 2 years interval	0.0447***
	(0.0101)
Order 5+ / 2 years or more interval	-0.0242***
	(0.00806)
Order 5+ / less than 2 years interval	0.0627***
	(0.0126)
Multiple birth	0.187***
	(0.0161)
HIV-positive mother	0.0927***
	(0.0116)
Year (2002 base category)	
2003	-0.0245***
	(0.00892)
2004	-0.0330***
	(0.00880)
2005	-0.0407***
	(0.00874)
2006	-0.0320***
	(0.00889)
2007	-0.0422***
	(0.00847)
2008	-0.0394***
	(0.00849)
Constant	0.104***
	(0.0139)
Observations	17,484
R-squared	0.043

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

Table B.IV 1st stage regression to instrument aid presence with the urban variable.

DEPENDENT VARIABLE	Aid presence, no lag
Urban	0.365***
	(0.0146)
Maternal education (no education base category)	
Primary	0.0117**
	(0.00552)
Secondary and higher	0.0433***
	(0.00994)
Maternal age at birth (less than 20 years base category)	
20-less than 35 years	0.00555
	(0.00733)
35 years or more	0.00303
	(0.0107)
Sex of child (male base category)	
Female	-0.00691
	(0.00440)
Birth order and interval between births (first order base category)	
Order 2-4 / 2 years or more interval	0.00686
	(0.00781)
Order 2-4 / less than 2 years interval	-0.00358
	(0.00937)
Order 5+ / 2 years or more interval	-0.00564
	(0.00921)
Order 5+ / less than 2 years inteval	-0.000211
	(0.0116)
Multiple birth	0.00320
	(0.0111)
HIV-positive mother	-0.000145
	(0.0105)
Year (2002 base category)	
2003	0.0172***
	(0.00530)
2004	0.0153***
	(0.00523)
2005	0.0799***
	(0.00706)
2006	0.177***
	(0.00838)
2007	0.171***
	(0.00777)
2008	0.169***
	(0.00765)
Constant	-0.0281**
	(0.0139)
Observations	17,484
R-squared	0.241

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

Table B-V 2nd stage IV regression. Aid presence specification: 10-km radius, binary, no lag. Aid presence is instrumented with urban.

DEPENDENT VARIABLE	Infant death
Aid presence, no lag	0.00693
	(0.0243)
Maternal education (no education base category)	
Primary	0.0109**
	(0.00530)
Secondary and higher	-0.00258
	(0.00811)
Maternal age at birth (less than 20 years base category)	
20-less than 35 years	-0.0168**
	(0.00686)
35 years or more	-0.00555
	(0.00977)
Sex of child (male base category)	
Female	-0.0121***
	(0.00386)
Birth order and interval between births (first order base category)	
Order 2-4 / 2 years or more interval	-0.0215***
	(0.00680)
Order 2-4 / less than 2 years interval	0.0447***
	(0.0100)
Order 5+ / 2 years or more interval	-0.0242***
	(0.00806)
Order 5+ / less than 2 years interval	0.0628***
	(0.0126)
Multiple birth	0.187***
	(0.0161)
HIV-positive mother	0.0929***
	(0.0116)
Year (2002 base category)	
2003	-0.0245***
	(0.00891)
2004	-0.0330***
	(0.00879)
2005	-0.0406***
	(0.00888)
2006	-0.0310***
	(0.00960)
2007	-0.0413***
	(0.00922)
2008	-0.0376***
	(0.00921)
Constant	0.103***
	(0.0139)
Observations	17,484
R-squared	0.043

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

Table B.VI List of all variables used in the propensity score matching.

Variable	Description
Malaria	Malaria endemicity categorized into low, high and very high risk
Education	Categories: No, primary and secondary education
Districts	The 27 administrative districts of Malawi
Year	Year of the observed birth
Wealth	A wealth indicator given by the DHS (5 categories)
Literacy	Assessment of literacy in five categories
Urban	Observation is within 10 km of an urban cluster
Ethnicity	Ethnicity of the respondent
Religion	Religion of the respondent
No of HH members	Number of people living in the household of the respondent
Age of HH head	The age of household head
Sex of HH head	Sex of household head
Husband	Husband stays in household
Other wives	Number of other wives in household
Daughters	Number of daughters living in the household
Electricity	Dummy variable for electricity in the household of the respondent
TV	Dummy variable for TV in the household of the respondent
Radio	Four categories for radio listening of the respondent
Watching TV	Four categories for TV watching of the respondent
Iodine	Test of salt in the household of respondent for iodine content
Distance to HF	Dummy variable equal to one, if it is a problem for the respondent to travel to the next health facility
Permission	Dummy variable equal to one, if it is a problem for the respondent to obtain permission to go to a health facility

Table B.VII Propensity score estimation, aid presence specification: binary, no lag and 10 km.

DEPENDENT VARIABLE	Aid presence, no lag
Maternal education (no education base category)	
Primary	0.0325
	(0.130)
Secondary and higher	-0.0564
	(0.181)
Malaria Endemicity (low risk is base category)	
High risk	0.0546
	(0.128)
Very high risk	-0.453**
	(0.185)
Wealth Index (poorest quartile is base category)	
Poorer	0.321**
	(0.129)
Middle	0.298**
	(0.129)
Richer	0.564***
	(0.135)
Richest	1.101***
	(0.168)
Close to urban area	4.542***
	(0.106)
Has electricity	0.467**
	(0.185)
Literacy (cannot read is base category)	
Able to read parts of sentence	0.179
	(0.145)
Able to read full sentence	0.458***
	(0.107)
Observations	16,439

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note, not all estimated coefficients are presented for clarity.