HEDG
HEALTH, ECONOMETRICS AND DATA GROUP
THE UNIVERSITY of York

WP 15/20

Free primary care in Zambia: an impact evaluation using a pooled synthetic control method

Aurélia Lépine, Mylène Lagarde & Alexis Le Nestour

September 2015

http://www.york.ac.uk/economics/postgrad/herc/hedg/wps/

Free primary care in Zambia: an impact evaluation using a pooled synthetic control method

Aurélia Lépine*
Mylène Lagarde
Alexis Le Nestour

June 2015

Abstract

We estimate the impacts of user fee removal in Zambia using a pooled synthetic control method. We find no evidence that user fee removal changed health seeking behaviours, even among the poorest. We show that these results are not attributable to the imperfect implementation of the policy. Nonetheless, our results confirm that the policy virtually eliminated medical expenditures, thereby providing financial protection to health services users. Since the poorest individuals were found to be less likely to use care and had lower expenses, ceteris paribus, the policy effect was similar to a transfer of US$2.22 per medical visit for the total sample but of only US$0.65 for the poorest people.

Keywords: user fees, health care use, health expenditure, provider choice, synthetic control, propensity score matching, difference-in-differences, Zambia.

JEL: C01, C20, I18

*Corresponding author: Aurelia.Lepine@lshtm.ac.uk, Department of Global Health and Development, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK. +00442079272845.
We are grateful to Adrian Gheorghe, Collins Chansa, Steven Stillman, Kelvin Seah, Seiro Ito, Noemi Kreif, Frank Kakungu, Richard Mutemwa, Tim Powell-Jackson, Richard Grieve and Neil Hawckins for helpful comments. The work was funded by the Medical Research Council UK and the LSHTM.

1 Introduction

The debate over whether low- and middle-income countries should charge their populations for using health care services has been highly contentious for several decades. Proponents of the introduction of user fees in the late 1980s have argued that such charges could play a positive role in extremely underfunded health systems (De Ferranti 1985, Akin, Birdsall et al. 1987), by providing additional resources to reduce drug shortages and top up health workers' salaries. Meanwhile, opponents of cost recovery policies have highlighted that a higher price-elasticity of demand for health care of the poorest groups meant that user fees would create insurmountable financial barriers to access the health services they needed, thereby fuelling inequity in health systems (Gertler, Locay et al. 1987, Creese 1991, Lavy and Quigley 1993, Sauerborn, Nougtara et al. 1994, Gilson, Russell et al. 1995, Gilson 1997). Following the examples of South Africa (Wilkinson, Gouws et al. 2001) and Uganda (Burnham, Pariyo et al. 2004, Nabyonga, Desmet et al. 2005, Xu, Evans et al. 2006), several African countries have taken steps to removing user charges partially or completely in the past decade (Yates 2009), embracing the emerging consensus that user fees "*deter people from using health services and cause financial stress*" (World Health Organisation 2010).

Despite the heated nature of this debate and its high policy relevance, there is very little rigorous evidence on the impact of making health services free by introducing a health insurance scheme or removing user charges at the point of use. A systematic review of the literature looking at the impact of removing or introducing user fees could only identify 16 studies that met the inclusion criteria (Lagarde and Palmer 2008). Amongst these, only five evaluated the impact of removing user fees in Madagascar (Fafchamps and Minten 2007), South Africa (Wilkinson, Gouws et al. 2001) and Uganda (Burnham, Pariyo et al. 2004, Nabyonga, Desmet et al. 2005, Xu, Evans et al. 2006) and concluded that the policy change had positive effects on health care utilisation and health-seeking behaviours. However, the evidence provided by these studies was deemed of poor quality for several reasons. First, due to the national implementation of these policy changes, none of these studies were able to compare changes in outcomes to a counterfactual, limiting the ability to establish causality between the health financing reform and the observed changes. Second, the majority of studies relied on poor quality utilisation data (facility registers) that contained limited patient-level information and were hampered by a large proportion of missing data and selection bias. In the absence of richer data, it was neither possible to analyse the equity implications of fee removal for different socio-economic groups, nor to ascertain whether removing fees provided enhanced financial protection. Similar conclusion was made in a more recent review of the literature focusing on the impact of fee removal on maternal and child health services (Dzakpasu, Powell-Jackson et al. 2013).

Since this systematic review, a few experimental studies have provided more robust evidence on the positive effects of providing free health care products or services. In several randomised

experiments looking at the uptake of health care products[1] at various price levels (including free product), economists have found evidence that price is an important determinant of health care use in low-income settings (Miguel and Kremer 2004, Thornton 2008, Cohen and Dupas 2010). Yet, the narrow focus of these studies on specific products provides limited lessons to inform the potential effects of system-wide health financing reforms, whose implementation have complex repercussions on various components of the health system – such as quality of care through effects on health workers or availability of drugs and equipment. Three recent experimental studies on the impact of health insurance have provided more directly relevant evidence for the debate on user fees. Using data from a randomised trial set in one district in southern Ghana, where a health insurance product to providing access to free health services was offered to randomly chosen households, Ansah, Narh-Bana et al. (2009) and Powell-Jackson, Hanson et al. (2014) showed that free care moderately increased utilisation of formal primary care by shifting care seeking away from informal providers. Yet this had limited effects on health outcomes, only improving health for those with anaemia before the trial. Using the step-wedge randomised implementation of a social health insurance in Mexico, King, Gakidou et al. (2009) found that although there was no increase in the use of health care services by those who took up the insurance (experimental compliers), the health insurance drastically reduced out-of-pocket (OOP) expenses. Thornton, Hatt et al. (2010) found similar results in a randomised experiment looking at the take-up and impact of insurance in the informal sector in Nicaragua, where insurance coverage providing free care did not appear to increase utilisation of services, even though it decreased OOP expenditures.

Yet, while the evidence emerging from field experiments is more robust to establish causal inference between the policy change (introduction of free care) and the outcomes of interest than the studies reviewed in Lagarde and Palmer (2008), it also presents several limitations that prevent from informing the debate on user fees. First, as it is often the case with evidence emerging from experimental studies, there are questions about their external validity. In particular, it has been pointed out that randomised field experiments are often carefully implemented, monitored and supported by experts, which departs from the usually 'messier' environment in which governments are used to implementing national reforms. Local authorities or front-line workers will typically be more reluctant or slow to implement a new policy than to take part in a high-profile innovative research study. Second, amongst the studies cited, the policy interventions closest to user fee removal are in fact the introduction of some form of health insurance. Although from a theoretical perspective the two are equivalent (no direct cost for using health care services if the insurance provides a full third party reimbursement), in practice they are quite different interventions that are likely to have different behavioural effects on individuals. Some of these studies cannot entirely rule out the selection bias problem of insurance (i.e. those who use health insurance products are different from those who do not), and therefore it is unclear whether their conclusions can be generalised to the whole

[1] Bed nets (Cohen and Dupas, 2010), deworming drugs (Kremer and Miguel, 2007) or HIV diagnostic tests (Thornton, 2008).

population. Insurance schemes often limit the number and type of health care providers that can be chosen by members, in a way that user fee removal does not. And there is also evidence that even when they are insured, more disadvantaged groups are likely to claim their benefits and use health insurance less (Devadasan, Criel et al. 2007), while user fee removal does not present any administrative obstacle to anyone.

Therefore it is unclear to what extent the universal introduction of free health care services might trigger changes in health-seeking behaviours. The aim of the paper is to answer this question, and more specifically to estimate the causal effect of user fee removal on health-seeking behaviours in the general population, and explore heterogeneous effects on the poorest. To do so, we apply a pooled synthetic control method to evaluate the impact of the policy change that occurred in rural Zambian districts in 2006 on health care use, the type of health provider chosen, and the level of out-of-pocket (OOP) medical spending.

Our paper makes three significant contributions to the literature. First, to our knowledge, this is the first study to provide evidence of the causal impact of removing user fees on health-seeking behaviours and out-of-pocket expenditures, in the context of a real-world policy change in a low-income setting. In practice, we use the opportunity offered by the policy change that occurred in Zambia in 2006 where user charges were removed in only 54 of the 72 districts in the country. This natural experiment provides a unique opportunity to test the short-term effect of a complex policy change that is known to have wide-ranging systemic effects.
Second, this paper provides an innovative application of a novel econometric tool, the synthetic control method, designed to estimate the causal effect of a policy in settings where a single unit is exposed to an event or intervention (Abadie, Diamond et al. 2010). Here, not one but 54 districts are exposed to the health financing reform, while 18 are not. Using several waves of nationally representative household surveys, we construct a synthetic control for each of the 54 districts where free care was introduced by using the characteristics of the 'control' districts before the policy implementation. We derive the impact of the policy change at the district level by taking the difference in the outcomes of interest, after the policy change, between each district and its synthetic twin. Following Dube and Zipperer (2013) and Gobillon and Magnac (2014), we pooled the 54 estimated effects to obtain the average treatment on the treated (ATT) at the national level by computing the weighted average (by district size) of the estimated district effects. Third, in contrast to a method that would construct a single treated unit based on the average characteristics in the 54 districts, this method allows to explore the determinants of the estimated district level ATTs and to further investigate the role of poor compliance on the impacts of the policy.

The results of the paper shows no evidence that user fees removal policy in rural Zambia increased the probability of seeking care, nor that it led to changes in provider choice, not even for the poorest individuals. We further show that this result is not attributable to the poor

implementation of the policy at the time of the analysis, six months after the policy change, since we find similar results in districts where the policy was fully implemented. Nonetheless, our results confirm that where it was well implemented, the policy virtually eliminated medical expenditures, thereby providing financial protection to health services users. However, given that the poorest individuals were found to be less likely to use care when sick and have lower OOP health expenses before the policy implementation, the policy mostly benefited the richer groups. In fact, we estimate that, ceteris paribus, the policy effect was similar to a transfer per medical visit of US$2.22 in the total sample but only of US$0.65 for the poorest people. Our results are robust to an alternative method for pooling the district-level estimated effects, to an alternative specification of the pooled synthetic control and to a propensity score matching combined with a difference-in-differences (DD) analysis.

The remainder of the paper is organized as follows. In section 2, we describe the study setting. Section 3 describes the econometric method adopted and section 4 presents the data. Section 5 presents the results and section 6 tests the robustness of the results. Section 7 discusses the findings and conclusions are presented in Section 8.

2 User fee removal policy in Zambia

From 1964 to 1991, the Zambian government provided free health care for all. Then in 1991, in line with the movement instigated by the World Bank and the Bamako Initiative, the government introduced user fees with the triple objective to raise additional income to improve quality of services, increase ownership of health system through community participation and improve staff motivation and accountability through salary top-ups (Government of the Republic of Zambia 1991). User fees in primary care consisted in a flat consultation fee set by the district that was based on the willingness to pay of the population (Carasso, Palmer et al. 2010). Children under 5 years old and adults over 65, certain diseases (e.g. HIV/AIDS, TB), pregnant women and indigents were officially exempted from paying user fees, and all referral services to higher level hospitals were supposedly provided free of charge when a referral letter was presented by the patient (Cheelo et al., 2010).

In January 2006, the Zambian president announced that user fees in health care were to be removed. The policy change was justified on the grounds given that user fees appeared to negatively affect equity of access to health care and contributed to the increase of poverty (Masiye, Seshamani et al. 2005). The policy was to be officially implemented three months later (starting on 1^{st} April 2006) in all primary health care facilities, health centres and district hospitals in the 54 districts designated as "rural" according to the local government classification. This decision by the Government of Zambia was a bid to make health services more affordable and accessible to the poor, particularly given that most of Zambia's poor reside in rural areas. In

January 2007, user fees were removed in all public health facilities located in the peri-urban areas of the 18 districts where fees had not been abolished.[2] Finally, in January 2012, free care was officially provided in the whole country. This staggered implementation of the policy change provides a unique opportunity to analyse the effects in the context of a quasi-natural experiment. Specifically, with 54 districts implementing the policy change in April 2006 while all public facilities were still charging user fees in the remaining 18 urban districts, we are able to clearly define treatment and control units. This is not possible for the second stage of the policy implementation, due to the subjectivity of the classification of peri-urban areas and the lack of detailed geographical information in the databases. This study therefore focuses on the evaluation of the 2006 health financing reform.

3 Method

3.1 Imperfect implementation of the policy change

As highlighted in the introduction, a key difference between a well-designed experimental study and a natural experiment stems from the implementation challenges typically faced in the real world when a policy change is announced and has to be translated on the ground. In effect, full and meticulous implementation rarely happens immediately, if at all. Accounting for the "good implementation" of the policy, i.e. the extent to which health care is actually offered for free to patients, is therefore central to understanding what effect is being measured.

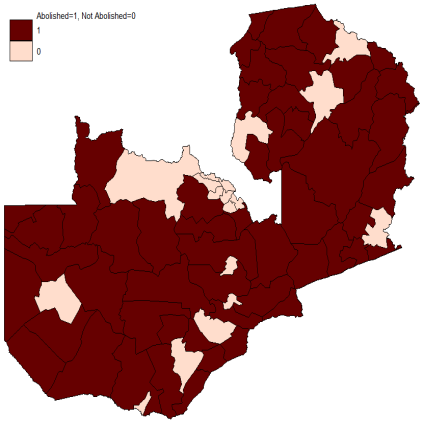
In order to measure the degree of implementation of the user fee removal policy in Zambia, we consider the proportion of patients in a rural district aged between 5 and 65 years[3] who sought care from a public health facility and declared to have received free care in the household survey. Map 1A (on the left hand side) in Figure 1 shows the rural districts that should have implemented the policy in dark. However, we can see from map 1B that the degree of implementation varies a lot: in several treated districts, the proportion of persons declaring having free care is below 50% and in some (rare) cases below 25%. On the contrary, in two of the control districts (Kasama and Mongu, indicated with a black dot), more than half of the respondents declared having received free care. This suggests contamination effect of the policy change from neighbouring (control) districts if one assumes that individuals of the urban districts that are surrounded by rural districts, may decide to seek care in a neighbour rural district to benefit from free care.

[2]Public health facilities outside the radius of 20 km of all the administrative centre of municipals and cities were exempted, but the absence of geographic coordinates in the data prevent us from measuring medium term effect of the policy.

[3]We remove from the analysis children under-5 and elderly given that they were already exempted before the policy implementation.

Figure 1: Implementation of the policy at the district level according to the policy (left) and in reality (right)

Map 1A: user fee abolition according to the policy



Map 1B: user fee abolition in practice

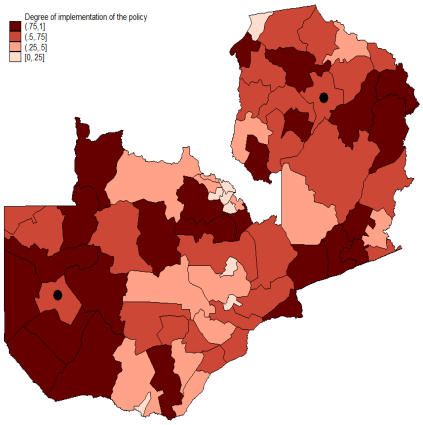


Table 1 describes the proportion of individuals receiving free care in the treated (rural) and control (urban) districts before (2004) and about six months after the policy implementation (2006). The table shows that six months after the official introduction of the free care policy, only 71% of individuals aged between 5 and 65 years who visited a public health facility in one of the 54 treated districts had actually received free care. In other words, 29% of patients in the treatment group were still paying for their OOP health expenditures.

Table 1: Proportion of individuals aged between 5 and 65 years receiving free care in public health facilities

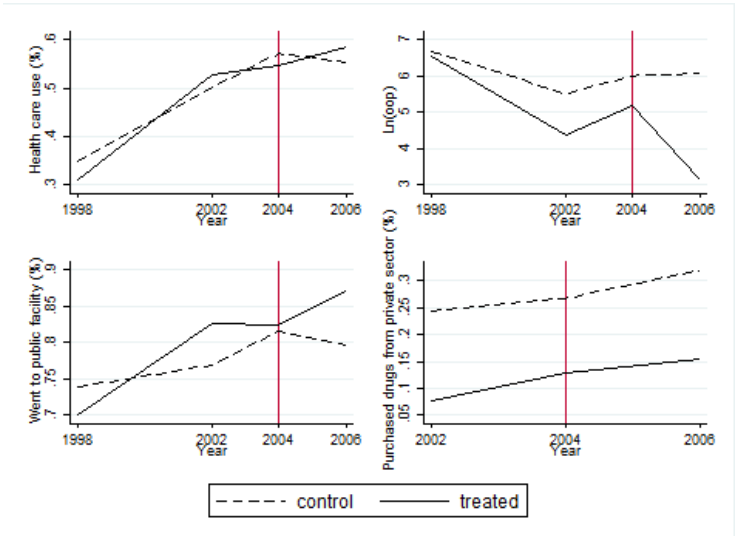
	In the 54 'treatment' districts	In the 18 'control' districts
Before (2004 LCMS data)	25.5%	32.4%
After (2006 LCMS data)	71.1%	30.6%
T-test, p-value	-32.7, p<0.01	-2.99, p<0.01

3.2 Synthetic control method

Since we have information before (1998, 2002, 2004) and after (2006) the policy change and on a sample of treated and control districts, we first test the parallel trend hypothesis i.e. whether the pre-intervention outcome trends were parallel in the control and treatment groups. This assumption is required in order to assess the effects of the policy by using a simple DD analysis. As suggested by Figure 2, we find that this fundamental assumption of the DD method is not

met for all outcomes of interest. Further testing the assumption of parallel trend at the district level by controlling for district fixed effects, we found that the assumption of parallel trend was only valid for health care utilisation and the demand for care from public health facilities.[4]

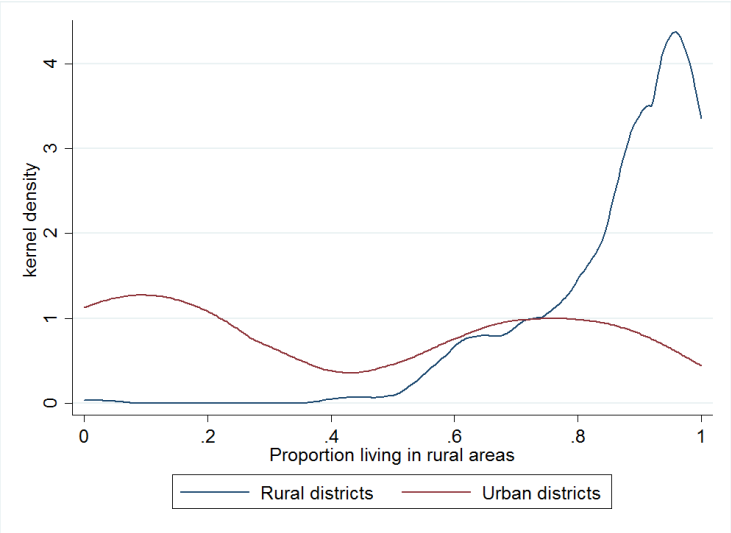
Figure 2: Trends of the outcomes of interest



Despite the fact that trends were not parallel, one may assume that a matching procedure would allow to reconstruct a group that share similar observed characteristics with the rural districts group. In fact, district classification was highly arbitrary since districts containing a city or a municipality were classified urban, while the other districts were classified rural. Nevertheless, looking at the distribution of the proportion living in rural areas in Figure 3, one can note that on average, 40 % of households living in a urban district live in a rural area and that there exist some highly rural districts among the urban classified districts.

[4]These results are available from the authors upon request.

Figure 3: Proportion living in rural areas in urban and rural districts



Therefore, in order to exploit the fact that some of the urban districts and rural districts may share some close characteristics, we evaluate the effects of the user free removal policy by using the synthetic control method. The synthetic control method involves constructing a control group based on a data driven procedure. For each treated unit, a synthetic control is created by taking a weighted average of the available control units, where a higher weight is given to control units that are more similar to the treated unit. The synthetic twin is created to follow the same pattern than the treated unit in the pre-treatment period so that it can be used as a counterfactual after the policy implementation. The use of the synthetic control methods over other quasi-experimental approaches is justified by various reasons. Unlike matching estimators, the idea behind the synthetic control is that a combination of control units provides a better comparison for the treated unit than a single unit alone. Additionally, the synthetic control is built by using the observable characteristics in all the pre-treatment years. In consequence, unlike the DD approach, the synthetic control method allows the effects of the unobserved heterogeneity in the outcome to vary with time.

Below, we describe the synthetic control method following the notation in Abadie, Diamond et al. (2010). Suppose that there are $J+1$ districts and T time periods $t=1,\ldots,T$, with an intervention I occurring in only one district, noted $i=1$ while all other districts i remain unexposed, these unexposed districts constitute the donor pool.

Let Y_{it}^{N} be the outcome observed for any district i at time t in the absence of the intervention (N stands for no intervention) and let Y_{it}^{I} be the outcome observed for the district exposed to the intervention. Let T_0 be the number of pre-treatment periods where $1=T_0<T$. For post treatment periods $t> T_0$, the outcome in the treated district in the absence of the intervention Y_{1t}^{N} is not observed.

Let $a_{it} = Y_{it}^I - Y_{it}^N$ be the effect of the intervention i.e. the difference in the outcome in the treated unit and the control units. Let D_{it} be an indicator that takes the value of 1 if unit i is exposed to the intervention at time t. The observed outcome is then given by a sum of untreated outcomes Y_{it}^N and the effect of the treatment a_{it} for district i and period t:

$$Y_{it} = Y_{it}^N + a_{it}D_{it} \tag{1}$$

For post treatment periods, $a_{it} = Y_{1t}^I - Y_{1t}^N = Y_{1t} - Y_{1t}^N$.
where the unobserved outcome in treated district in the absence of the intervention Y_{1t}^N needs to be estimated.
Suppose that Y_{it}^N is given by a factor model:

$$Y_{it}^N = \delta_t + \omega_t Z_i + \lambda_t \mu_i + \epsilon_{it} \tag{2}$$

where δ_t is a time fixed effect, ω_t is a vector of parameters associated to the vector of time-invariant observed covariates Z_i, λ_t is a vector of parameters associated to the vector of time-variant unobserved covariates μ_i and error terms ϵ_{it} are unobserved transitory shocks at the district level with zero mean. Note that it is because μ_t is not constant for all t that the synthetic control approach differs from the traditional DD and allows the effects of confounding unobserved characteristics to vary with time.
To estimate the unobserved outcome in treated district in the absence of the intervention Y_{1t}^N, a weighted average of the potential control units *i=2,...J* needs to be estimated. To do this, let $W = (w_2, \ldots, w_{j+1})'$ be a vector of weights such that $w_j \geq 0$ for *J=2, ..., J+1* and $w_2 + \cdots + w_{j+1} = 1$.
Let denote X_1 a vector of pre-intervention observable characteristics for the treated districts and X_0 a matrix that contains the same variables for the control districts. The optimal vector W^* is chosen to minimise the distance between X_1 and X_0W. To measure the discrepancy between X_1 and X_0W:

$$\|X_1 - X_0W\| v = \sqrt{X_1 - X_0W)'V(X_1 - X_0W)} \tag{3}$$

where V is a symmetric and positive semi-definite matrix that minimises the root mean squared prediction error (RMSPE) i.e. the average of the squared discrepancy between the level of outcomes in the treated unit and in its synthetic control counterpart in the pre-treatment periods.
Using these weights, the synthetic control is given by:

$$\hat{Y}_t^N = \sum_{j=2}^{J+1} w_j^* Y_{jt} \tag{4}$$

And for the post-treatment periods $t > T_0$, the effect of the treatment for each treated district

is:[5]

$$\hat{a}_{1t=}\ Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt} \tag{5}$$

Unlike the seminal case presented by Abadie, Diamond et al. (2010) where a single treated unit is considered, here user fees were removed in all the treated districts *i=1,...,*L_1 where L_1=54 treated units, while districts remaining in the donor pool are l = L_1+*1,...,* L_1+ L_2 where $L_2 = 18$:

$$\hat{Y}_{it}^N = \sum_{l=L_1+1}^{L_1+\ L_2} w_l^* Y_{lt} \tag{6}$$

This implies that we obtain $\hat{a}_{it}$ for each treated unit *i=1,...,*L_1 :

$$\hat{a}_{it} = Y_{it} - \hat{Y}_{it}^N \tag{7}$$

3.3 Pooling the treatment effects

To pool the treatment effects obtained from each treated district, we follow the approach described by Dube and Zipperer (2013) and simply take the average of the estimated effects $\hat{a}_{it}$ weighted by the district size P_i:

$$\overline{a}_t = \frac{\sum_{i=1}^{L_1} (\hat{a}_{it} P_i)}{\sum_{i=1}^{L_1} P_i} \tag{8}$$

where $\overline{a}_t$ denotes the average treatment effect on the treated of the policy and P_i is defined as the population of district i. Similar method was used to estimate $\overline{RMSPE}_{it}$.

The statistical significance of $\overline{a}_t$ is estimated by following the method of Abadie, Diamond et al. (2010) that consists of randomly permuting the treatment status in the pool of donor districts in order to construct a synthetic control for each control district. The idea is that the policy should not have affected the control units, thus for the pre-treatment periods, the level of outcomes in the averaged treated unit should follow a close trend to the trends of units in the donor pool. However, in the post–treatment period, the trend of the treated unit should significantly deviate from the trends of the placebo group. To assess this significance, we estimate the 90% confidence interval from the distribution of the placebo estimates in the post treatment year by using the percentile rank statistics p. This method consists in estimating whether the effect in the averaged treated unit $\overline{a}_t$ lies in the tail of the resulting placebo distribution formed by $\hat{a}_{it}$ where l = L_1+*1,...,*L_1+ L_2. We reject the null hypothesis $\overline{a}_t = 0$ when $\hat{a}_{it}$ is not included in the 90% confidence interval, i.e. when $p(0.05) > \hat{a}_{it} > p(0.95)$.
The contamination of two districts from the donor pool (presented in Figure 1) violates a

[5]Note that regarding the implementation, the nested optimisation was ran three times using three different starting points as a robustness check.

key assumption of the synthetic control method. The fact that control districts have had a higher degree of implementation than treated districts could bias the estimated effects as well as the placebo tests. For these reasons, following Abadie, Diamond et al. (2010), the two contaminated districts were removed from the donor pool and are not considered in this study.

3.4 Exploring the determinants of treatment effects heterogeneity

Given that this method of applying synthetic control with multiple treatment units results in the estimation of 54 district-level treatment effects, it is then possible to investigate if the effect $\hat{a}_i$ in the post-treatment period is influenced by observable characteristics of the district, such as the degree of implementation of the policy.

$$\hat{a}_i = I_i + V_i + U_i \tag{9}$$

where $\hat{a}_i$ is the effect of the policy in district i, I_i is the degree of implementation of the policy measured by the proportion of eligible individuals who received free care, V is a set of observable characteristics of district i and U is an error term.

4 Data

The estimation strategy exploits the Living Conditions and Monitoring surveys (LCMS) collected in 1998, 2002/2003, 2004 (pre-treatment) and 2006 (post-treatment) since data collection for the 2006 LCMS took place in October-November 2006, or six months after the policy change. We could not use data from survey taking place before 1998 due to the fact that only 54 districts were surveyed out of the 72 districts. In fact, there was an administrative reform in 1991 that lead to the creation of 18 new districts but these districts were only included in the national surveys from 1998. There was also a major change in the way questions relating to health-seeking behaviours were asked in the 1998 survey and the subsequent waves.

The LCMS is a nationally representative household survey, providing detailed information about health-seeking behaviours, as well as a variety of socio-economic variables on 327,616 individuals between 1998 and 2006. A similar sampling procedure was used for all surveys. Districts were divided into wards, subdivided into census supervisory areas, subdivided into standard enumeration areas (SEAs) that constituted the primary sampling units. The survey employed a two-stage stratified cluster sample design where first, SEAs were selected with probability proportional to estimated size, then households were selected from each selected SEA. We use sampling weights in analyses to account for this sampling procedure.

Since the user fees removal policy was implemented at the district level and because ward/census area and/SEA level analysis was not possible to due sampling reason, we evaluate the policy at this level, and, using survey sampling weights, we collapse all outcomes and independent

variables at the district level. We obtain a balanced panel of 72 districts observed over three pre-treatment periods (1998, 2002/2003, 2004) and one post-treatment period (2006). We implement the synthetic control method to investigate the effect of the policy on four distinct outcomes:

1. Health care utilisation defined as the likelihood of seeking care based for individuals who have experienced an illness episode the last 2 weeks
2. Provider choice defined as the likelihood of going to a public health facility
3. The likelihood of buying drugs from a private pharmacist to detect a potential unintended negative effect of the policy change – whereby increased utilisation of public health facilities might lead to drug stock-outs and the necessity to buy drug from private outlets
4. Health care expenditures defined as the amount of out-of-pocket (OOP) medical expenses incurred for those who sought care the last 2 weeks. OOP medical spending is deflated and expressed in Kwachas 2006.

The last three outcomes are conditional on using care.

Since they already benefited from free care before 2006, at least in theory, we exclude from the analysis individuals aged less than 5 and more than 65 years old. Indeed, detailed information in the 1998 survey on exemption reasons indicate that these represent the majority of cases, specifically 69% and 6% of individuals who declared to have benefited from free care were respectively children under 5 year old and elderly.

We construct a synthetic control for each of the 54 treated districts by using 16 control districts (we exclude the two control districts where we suspect that contamination happened as a high proportion of individuals benefited from free care – see section 3.1). To optimise the distance matrix and create the synthetic controls, we include the following district level covariates in addition to the pre-intervention outcomes levels: proportion of male, proportion of female heads of households, proportion of households living in a rural area, median age, median household income, median household size, median distance to the health facility. Note that those covariates were chosen to be included because we find that the RMSPE is minimised under this model. The same covariates were used for all the outcomes except for the likelihood of buying drugs from the private sector where information for 1998 was not contained in the data.

The analysis was carried out in STATA using the -synth- package.

5 Results

5.1 Descriptive statistics

The data indicate that about 10% of the sample in 2004 declared to have been sick two weeks prior to the survey. Most of the types of illness reported were malaria (39%), cough (13%), diarrhoea (7%), backache (6%) and skin infection (3%). On average, 58% sought a treatment in 2004 while 61% sought a treatment in 2006 after the policy implementation. In 2004, 82% of those who sought a treatment went to a public provider but among these, 4% bought drugs from the private provider. In total in the sample, 19% purchased drugs from a private drug seller. Health expenditures were log-transformed given their skewed nature and were about 11,992 (US$ 3.15) 2006 Kwatchas in 2004. The average distance to the closest facility is 4.3 kilometres. Regarding socio-economic characteristics, on average individuals are 21 years old and belong to household of 7.5 members. The average household monthly income proxied by household monthly expenditure is 947,175 current 2006 Kwatchas (US$ 249). The level of education is low (20% has never attended school) but head of the households are educated since only 9% have never attended school. On average 52% of households live in a rural area in the sample.

Before estimating the impact of the policy, it is important to question about the characteristics of health care users before the policy implementation. Indeed, an analysis of the correlates of health-seeking behaviour may provide some evidence about the role that income plays on health-seeking behaviours. The analysis of the determinants of the outcomes of interest (health care use, health expenditure and provider choice) at the individual level in the pre-baseline data sets (1998, 2002, 2004) is presented in Table 2. We find that the richest individuals are more likely to use care when sick, have a slightly lower likelihood of going to a public health facility and also a slightly higher likelihood of buying drug from the private sector. Richer individuals have higher OOP medical expenses. However, if we look at the coefficient associated with income, we can see that an increase in one point in the logarithm of total expenditures (which is similar in the data to an increase in one tercile in the income distribution) increases health care use by only 4.5 percentage points. This result highlights the absence of a strong financial barrier to health care use. One may want to note that the coefficient associated with income could be endogenous to health care use due to the omission of unobserved individual (preferences for health, cultural factors, etc.) and community (quality of care) characteristics. However these variables are likely to be correlated with income and health care use in the same direction, so it is unlikely that the effect of income obtained is underestimated.

Table 2: Determinants of the outcomes analysed in the pre-intervention period

	(1) Health care use		(2) Use of a public facility		(3) Use of a private pharmacist		(4) Out-of-Pocket expenditures (ln)	
VARIABLES	Coef	SE	Coef	SE	Coef	SE	Coef	SE
Household size	-0.003	(0.002)	0.004*	(0.002)	0.003**	(0.001)	0.018	(0.017)
Age	-0.000	(0.001)	-0.002	(0.001)	-0.000	(0.001)	0.024**	(0.012)
Age^2	0.000	(0.000)	0.000	(0.000)	-0.000	(0.000)	-0.001***	(0.000)
Distance to facility	-0.006***	(0.001)	-0.004***	(0.001)	0.001*	(0.000)	0.010	(0.007)
Rural	0.006	(0.013)	-0.024*	(0.013)	-0.115***	(0.010)	-0.723***	(0.112)
2002 (ref:1998)	0.017	(0.089)	-0.122*	(0.069)			-0.959*	(0.574)
2004	0.054	(0.088)	-0.105	(0.068)	0.022**	(0.008)	-0.365	(0.568)
Male	0.033***	(0.011)	0.000	(0.011)	-0.002	(0.008)	-0.654***	(0.099)
Ln(expenditures)	0.045***	(0.007)	-0.025***	(0.007)	0.031***	(0.005)	0.533***	(0.061)
Head	-0.107***	(0.015)	-0.027*	(0.016)	0.055***	(0.011)	1.791***	(0.133)
Female head	-0.008	(0.017)	0.007	(0.017)	-0.020*	(0.011)	-0.158	(0.145)
District population	-0.000***	(0.000)	-0.000	(0.000)	0.000***	(0.000)	0.000**	(0.000)
Observations	16,210		9,389		13,068		13,166	
R-squared	0.040		0.024		0.121		0.093	

Estimates obtained through OLS regression, estimates include province dummies, sampling weights and are representative at the national level. Robust standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1

From the household data set, we computed means and medians of the relevant variables included in the distance matrix at the district level using survey sampling weights. The descriptive variables at the district level for the total sample per year are presented in Table 3.

Table 3: Descriptive statistics of independent and dependent variables by year at the district level

	1998 (n=72)		2002/2003 (n=72)		2004 (n=72)		2006 (n=72)	
Variables	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Health care use (%)	0.327	0.142	0.523	0.148	0.573	0.120	0.609	0.146
Went to public health facility (%)	0.739	0.226	0.804	0.196	0.808	0.143	0.862	0.109
Bought drugs from private sector (%)	_	_	0.116	0.120	0.150	0.133	0.170	0.128
Log of deflated OOP	6.468	1.609	4.684	1.584	5.277	1.163	3.556	1.747
Male (%)	0.494	0.025	0.488	0.025	0.494	0.017	0.486	0.021
Female head (%)	0.020	0.055	0.016	0.032	0.018	0.053	0.025	0.034
Ln(expenditure)	13.329	0.399	13.233	0.305	13.132	0.338	11.957	0.697
Median age	19.722	1.201	20.569	1.320	20.542	0.934	20.042	1.250
Median distance to facility	4.056	3.863	4.236	4.698	3.347	2.733	4.542	4.206
Median household size	5.875	0.854	9.125	1.433	6.847	1.122	8.639	1.079
Proportion of households living in rural area (%)	0.777	0.284	0.785	0.294	0.704	0.280	0.693	0.263
Implementation of the policy in 2006 (%)	_	_	_	_	_	_	0.630	0.229
Population in 2010	181,201	209,967	181,201	209,967	181,201	209,967	181,201	209,967

5.2 Effects of the policy

Table 4 presents the impacts of the policy[6] and the 90% confidence interval estimated from the placebo tests presented in Figure 4.[7] Average effects at the national level are estimated by taking the weighted average of the ATTs by the district population size.

The findings show that there is no evidence that the policy had an impact on health-seeking behaviours. In addition, there is also no evidence that the policy has changed provider choice, since we observe that the likelihood of seeking care from a public facility has only increased by 2.4 percentage points and the likelihood of buying drugs from the private sector only decreased by 5.8 percentage points, both results being not statistically significant. However, the results indicate that the policy reduced out-of-pocket medical expenditures by 2.224 logarithm points, which corresponds to a decrease in 89% of OOP expenditures[8] compared to 2004 level. This represents a saving during the last medical contact of USD$2.22 (in 2006 USD) [9] or 0.9% of monthly household income, ceteris paribus, i.e. assuming that the level of consumption of

[6]The graphs for the 4 outcomes that illustrate the quality of the pre-intervention fit of the synthetic control are presented in Appendices 1-4.

[7]Note that sampling weights were also included for the estimation of the trends in the donor pool.

[8]1-(exp(-2.224))=0.892

[9]OOPs level in 2004 was Kwatchas 9480 or 2006 USD$2.49

health services remained unchanged in 2006.

Table 4: Impacts of user fee removal policy in rural Zambia in 2006, estimated on the whole sample

		Placebo test	
Outcomes	Estimated effect	5^{th} percentile	95^{th} percentile
Health care use	-0.002	-0.144	0.122
Public health facility	0.024	-0.259	0.208
Private drug seller	-0.058	-0.237	0.196
Out-of-Pocket Expenditures (ln)	-2.224	-1.001	1.558

Figure 4: Placebo tests

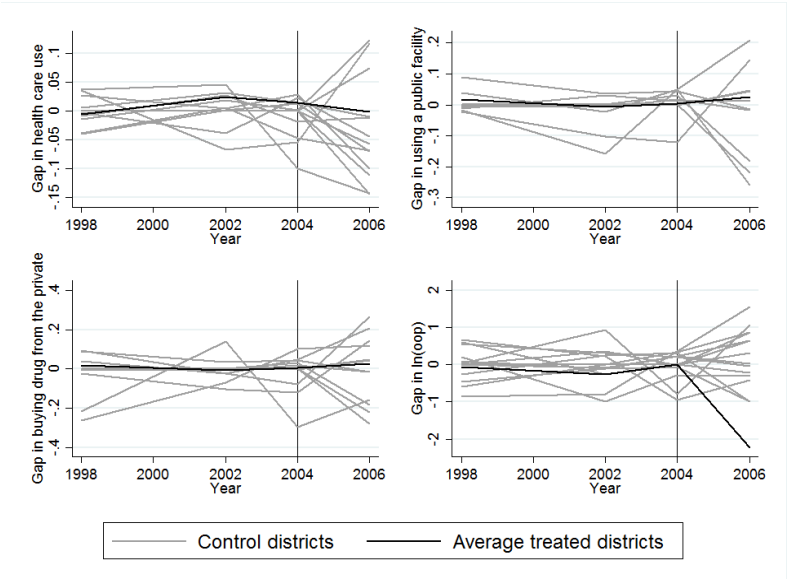


Table 5 presents the same results, but estimated on the poorest tercile of the population.[10] We find that the effect on OOP medical expenses previously mentioned is much lower as OOP expenses only decrease by 67%[11], which corresponds to a saving of US$0.65[12] and this result was not statistically significant. Additionally, we find that while the effect on health care use is greater, the effect is still not statistically significant. The impacts of the policy on the choice of the health facility and drug seller remain the same, i.e. not statistically significant. These results then do not show any evidence of a change in health-seeking behaviours among the

[10]We use the tercile, since we experienced some convergence issues when using smaller groups of the distribution (quintile, quartile).

[11]1-(exp(-1.115))=0.672

[12]OOPs level in 2004 was Kwatchas 3715 for the poorest tercile or 2006 USD$0.97

poorest individuals.

Based on the data, we estimate that about 12% of the total population (and this proportion is the same across the poorest and richest terciles[13]) will seek care over a 1-month period in the last pre-intervention period. Using the information on health expenses, a random person from the poorest tercile will spend on average US$0.095 over the last month while an average individual will spend US$0.36. Using the reduction in OOP expenses that resulted from the implementation of the policy we estimate that the general population has benefited from a monthly transfer 5 times greater than the one the poorest individuals received.[14] This result relies on the hypothesis that the likelihoods of being sick and of seeking care have not been altered by the policy, which is consistent with our results. It also relies on the assumption that the consumption of services has been unchanged by the policy, which is something we cannot test with the data.

Table 5: Impacts of user fee removal policy in rural Zambia in 2006, estimated on the poorest tercile

		Placebo test	
Outcomes	Estimated effect	5^{th} percentile	95^{th} percentile
Health care use	0.058	-0.214	0.306
Public health facility	0.038	-0.246	0.142
Private drug seller	-0.056	-0.204	0.134
Out-of-Pocket Expenditures (ln)	-1.115	-1.028	2.864

5.3 Factors affecting the effects of the policy in the treated districts

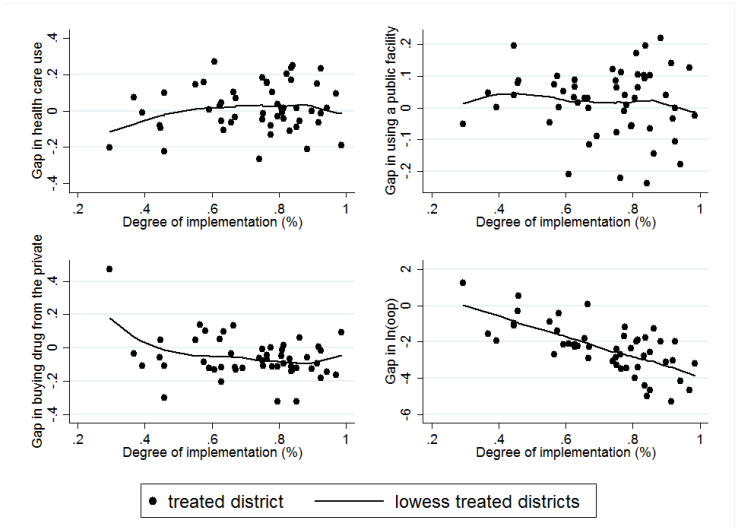
The ATT at the national level was computed by averaging out the estimated effects in the 54 treated districts. In the previous section, we presented the absence of evidence of an effect of the policy at the national level on health-seeking behaviours. Nonetheless, one may want to argue that this finding could be attributable to the poor implementation of the policy, highlighted in Figure 1. Note that the same definition as the one presented in Figure 1 was used to measure the quality of the policy implementation i.e. the proportion of patients in a rural district aged between 5 and 65 years who sought care from a public health facility and declared to have received free care. In order to answer this question, we further investigate the factors explaining the heterogeneity in the treatment effects estimated.

[13]Poor individuals are more likely to be sick but less likely to use care when sick, which leads to a similar proportion of individuals who use care among the richest and the poorest tercile.

[14]An average individual and a random individual from the poorest tercile will receive a monthly transfer of a value of US$0.32 and US$0.063 respectively.

Table 6 presents the determinants of the effect of the policy, estimated in the treated districts. Results presented in column (1) interestingly show that there is no significant relationship between the degree of implementation of the policy and its effect on health care utilisation, suggesting that the lack of impact of the policy is probably not due to its poor implementation. Therefore, unless there are unobservable factors associated with the degree of implementation of the policy and the heterogeneity of its effect on health care utilisation, this result casts doubt on the fact that the removal of user fees would have had a positive impact on health care utilisation, even if it had been fully implemented everywhere. In addition, as show columns (2) and (3), the poor implementation does not also explain why the policy had not affected the type of health provider visited. However, column (4) shows that the degree of implementation of the policy has a statistically significant impact on the reduction of OOP expenditures. We estimate the effect of the degree of implementation on the reduction in out-of-pocket expenses by looking at the reduction in OOP medical expenses in the district with the highest implementation of the policy (that is 97% in the data). As expected, we find that in this district, removing user fees leads to reduction in OOP by 99%.[15] Figure 5 shows a graphical representation of these findings.

Figure 5: Relationship between the impact of the policy and its degree of implementation



Overall, the results presented in Table 6 show that observable characteristics of the district do not play an important role on the effect of the policy estimated in the district. Nonetheless, we can note that the estimated effects, especially the effect on utilisation, decreases with the degree of ruralness of the district. This result may suggest that highly rural districts may share unobserved characteristics that negatively affected the impacts of the policy such as poor infrastructures and lower quality of care.

[15] 1-(exp-4.66)=0.99

Table 6: Determinants of the estimated impacts of the policy at the district level

	(1)		(2)		(3)		(4)	
	ATT of the policy on health care use		ATT of policy on the likelihood of using a public facility		ATT of the policy on the likelihood of using a private pharmacist		ATT of the policy on Out-of-Pocket Expenditures (ln)	
VARIABLES	Coef	SE	Coef	SE	Coef	SE	Coef	SE
Degree of implementation	-0.015	(0.136)	0.023	(0.106)	-0.317	(0.194)	-4.134***	(1.013)
Median distance to facility	-0.091	(0.100)	0.028	(0.081)	-0.135	(0.109)	-0.295	(0.990)
Median income	-0.051	(0.078)	-0.037	(0.057)	-0.010	(0.067)	0.266	(0.804)
Distance to facility*Income	0.008	(0.009)	-0.002	(0.007)	0.011	(0.010)	0.023	(0.087)
Median age	-0.043**	(0.018)	0.018	(0.012)	-0.011	(0.019)	0.223	(0.171)
Population density	-0.001	(0.001)	-0.002*	(0.001)	0.002	(0.001)	0.027*	(0.015)
Rural	-0.645***	(0.227)	-0.406**	(0.201)	0.013	(0.255)	3.129	(2.900)
Observations	51		53		53		51	
R-squared	0.294		0.556		0.412		0.615	

Standard errors in parentheses. Note that there were some convergence issues for some of the 54 districts, which explains the number of observations.*** p<0.01, ** p<0.05, * p<0.1. Estimates include province dummies.

6 Robustness checks

6.1 Alternative measures of estimating national level effect

Given the absence of methodological evidence regarding the most appropriate method to pool the estimated effects in each treated unit, we test the robustness of the results by using alternative methods to estimate the effect of the policy at the national level. In Table 7, we present three additional methods of pooling estimated effects to the one that weighted average by population size (column (1)). More specifically, column (2) presents the weighted median by the population size. Column (3) presents estimated pooled effects using the weighted average by the inverse of the RMSPE in order to give less weight to more imprecise estimates. In fact, graphs presented in Appendices 1-4 show that the pre-intervention goodness fit was highly heterogeneous among districts. This method of pooling district-level effects is then used as an alternative of trimming on pre-intervention goodness fit using the RMSPE. Finally, in column (4), we also present the aggregated effects by giving a higher weight to districts that better implemented the policy in order to account for results presented in the previous section. We find that different methods of pooling the effects at the district level do not affect our main findings.

Table 7: Different methods for estimating the national effect of the policy for the whole sample

	(1)	(2)	(3)	(4)
	Weighted average by population size	Weighted median by population size	Weighted average by the inverse of the RMSPE	Weighted average by the degree of implementation
Health care use	*0.003*	-0.002	-0.006	0.019
Public health facility	*0.024*	0.024	0.021	0.017
Private drug seller	*-0.047*	-0.058	0.020	-0.068
Out-of-Pocket Expenditures (ln)	*-2.184*	-2.223	-1.696	-2.562

6.2 Alternative pooling method

To estimate the national effect of the policy, we use a different approach to pooling, specifically the one used in Kreif, Grieve et al. (2014). Instead of pooling the effect estimated for every single treated unit *i=1,...,*L_1, we first create a unique treated unit $\widehat{Y_{1t}}$ by collapsing the outcomes from the 54 treated districts using the survey sampling weights S_{it}:

$$\widehat{Y_{1t}} = \frac{\sum_{i=1}^{L_1} Y_{it}\ S_{it}}{\sum_{i=1}^{L_1} S_{it}} \tag{10}$$

$$\hat{a}_{1t} = \widehat{Y_{1t}} - \frac{\sum_{l=L_1+1}^{L_1+\ L_2} w_l^*\ Y_{lt} S_{lt}}{\sum_{l=L_1+1}^{L_1+\ L_2} S_{lt}} \tag{11}$$

If $\widehat{Y_{1t}}$ refers to this averaged treated unit while all other units $l = L_1$*+1,...,* $L_1 + L_2$ *r*emain unexposed and constitute the donor pool, then the ATT $\hat{a}_{1t}$ is given by equation (11).
Having a single treated unit allows to plot the trend of the treated unit and the synthetic control on a single graph as shown in Figure 6. From this figure, the pre-treatment trend in the averaged treated unit and in its synthetic control perfectly overlap for the four outcomes analysed, which ensures that the trend of the synthetic control after the policy implementation can be used as a counterfactual. Outcome predictor means in the treated unit and in its synthetic control are presented in Appendix 5a and district weights in the synthetic control in Appendix 5b.

Figure 6: Trends in the averaged treated unit and in its synthetic control

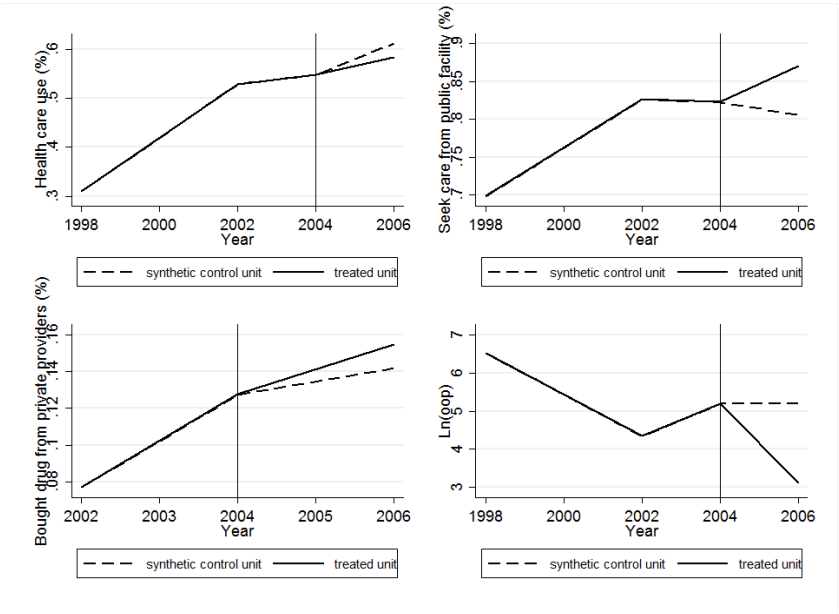


Table 8 presents the results obtained with this alternative method of estimating a pooled synthetic control. The conclusions are similar to the ones obtained by averaging out the estimated effects for the 54 treated districts. We still find that the policy has significantly reduced OOP expenditures but that there is no evidence of a change in health-seeking behaviours. The use of a single treated unit leads to close coefficients then when using 54 treated units. However, it is interesting to note in Table 9 that by using this method, the RMSPE is closer to 0. This is because a large share of the RMSPE when estimating a synthetic control for each treated unit is attributable to sampling errors. In fact, we find that there is negative correlation between the RMSPE and the district population [16] given that when estimating average from the whole sample, a larger number of observations is used. In addition, given that this method consists in estimating an average for the whole sample before constructing the synthetic control, the method reduces the existence of extreme values that may facilitate the matching procedure.

[16]The size of the sample per district is proportionate to the size of this district.

Table 8: Impact of user fee removal policy in rural Zambia in 2006 on the whole sample using a single treated unit

		Placebo test[17]	
Outcomes	Estimated effect	5^{th} percentile	95^{th} percentile
Health care use	0.013	-0.144	0.122
Public health facility	0.064	-0.259	0.208
Private drug seller	-0.027	-0.237	0.1962
Out-of-Pocket Expenditures (ln)	-2.078	-1.001	1.558

Table 9: RMSPE depending on the pooling method used

	54 treated units	1 treated unit
Health care use	0.057	2.64×10^{-13}
Public health facility	0.036	2.41×10^{-14}
Private drug seller	0.012	7.89×10^{-13}
Out-of-Pocket Expenditures (ln)	0.439	2.98×10^{-12}

6.3 Propensity score matching combined with difference-in-difference (DD)

As a last robustness check, we implement a propensity score matching combined to DD. Using propensity score matching, we recreate the control group for each treated district by exploiting propensity score for treated and control district i.e. the conditional probability of the district of receiving the policy, given district pre-treatment characteristics, this was done by using the last pre-treatment year (2004).[18] The pre-treatment characteristics selected were the ones that provided the greatest overlap (presented in Appendix 6) and that satisfy the test of balancing property of the propensity score (using a significance level of 0.01). The R-squared obtained in the Probit estimate of the likelihood of removing user fees is 0.55. The results from the Probit estimate indicate that rural districts are more likely to be treated. In treated districts, children under-5 are less likely to have a clinic card and to be immunised against DPT. Finally, households living in treated district are poorer. The propensity score was estimated by using a

[17]Placebo tests for this pooling method may be misleading as the sample sizes of control units are relatively small compared to sample size of the pooled treated unit and therefore errors will be larger.

[18]The pre-treatment characteristics were the proportion of population living in rural area, proportion of children under-5 with clinic card, prevalence of immunisation against DPT, median age, median distance to health facility, median income, proportion of civil servant, proportion of households with electricity access, proportion of households with access to drinking water, proportion of households with flush toilet. Results of the Probit are presented in Appendix 6.

blocking estimator (Cochran 1968, Rosenbaum and Rubin 1984). Extreme values were dropped following Crump, Hotz et al. (2008) and we imposed the common support restriction to improve the quality of the matches. We matched on the propensity score to create a balanced sample. We present the results using both the nearest neighbour and the kernel matching (Gaussian kernel) estimators. We matched with replacement i.e. that a control unit could be used as a match for more than one treated unit. We use sampling weights in order to account for the sampling strategy.

Given the nature of our data, we then combine the propensity score matching with DD (Heckman, Ichimura et al. 1998, Imbens 2004, Abadie 2005). Following Blundell and Dias (2002), the matching estimator combined with DD (MDD), noted a^{MDD} is given by:

$$a^{MDD} = \sum_{i \in T} \left\{ [y_{it_1-} \quad y_{it_0}] - \sum_{j \in C} \tilde{w}_{ij} [y_{jt_1-} \quad y_{jt_0}] \right\} w_i \tag{12}$$

where district i and j belongs to the treatment (T) and control (C) groups respectively, and y_{it_1} is the outcome of treated district i at time t while y_{it_0} refers to the outcome in non-treated districts. $\tilde{w}_{ij}$ is the weight placed on comparison observations j for treated district i and w_i accounts for the reweighting that reconstructs the outcome distribution for the treated sample.

Results obtained using the propensity score matching combined with DD give similar results to the ones obtained using the synthetic control method as shows Table 10. A difference is that for the choice of provider (health facility and pharmacist) the estimated coefficients are slightly larger and only the coefficient associated with the type of pharmacist is statistically significant at 10% with the kernel matched DD. The results obtained with matching combined to DD also indicate that the policy has significantly reduced the level of OOP. The effect of the policy on OOP medical expenses is estimated to be a reduction of about 91 percent, which is line with the results obtained with the synthetic control method.[19]

[19] 1- [exp(-2.443)]=0.913 and 1- [exp(-2.351)]=0.905

Table 10: Treatment effects on the treated estimated using matching combined with DD

	PS nearest neighbour matched DD		PS kernel matched DD	
	coeff	SE	coeff	SE
Health care use	-0.048	(0.077)	-0.023	(0.065)
Public provider	0.071	(0.097)	0.082	(0.06)
Private drug seller	-0.064	(0.09)	-0.076*	(0.048)
Out-of-Pocket Expenditures (ln)	-2.443***	(0.642)	-2.351***	(0.404)

*** p<0.01, ** p<0.05, * p<0.1. PS Kernel matched standard errors are obtained by bootstrapping using 1000 iterations, sampling weights are included using the population in the district in 2010. The estimates exclude the two control districts with a degree of implementation superior to 50% and only include the observation on support.

7 Discussion

In this paper, we assessed the short-term impacts of removing user charges for all primary and secondary care services provided by public health facilities in Zambia. Although this health financing policy has been the focus of one of the most heated policy debated in health economics, in particular in low- and middle-income countries, no study has previously assessed the causal impact of such reform undertaken on a large-scale. Using the fact that the policy was first implemented in rural districts only, we construct a synthetic control based on the characteristics of urban districts. We found that the policy change was successful in reducing overall out-of-pocket medical expenditures. However, we found no evidence that six months after its implementation, the policy change had led to an increase in the use of health services or in a change in the choice of health care providers.

Our main finding that there is no evidence of meaningful effect of user fee removal on health-seeking behaviours contradicts the results of previous past studies investigating the effect of user fee removal in Zambia (Masiye, Chitah et al. 2008, Onde 2009, Lagarde, Barroy et al. 2012). Those studies used health facility routine data and applied interrupted time series to investigate the effect on health care use and found that the increase in the number of visits recorded peaked 6 months after the implementation of the policy and that this effect flattens out over time. On average, they find an increase in utilisation of about 40% in the six months that followed the policy implementation. We have several reasons to think that the results reported in these studies may suffer from several weaknesses that may explain the discrepancies between the results they reported and the ones we present. Firstly, facility registers data may suffer from strong measurement errors especially given the importance of missing data. Secondly, routine data were only collected by some facilities of some districts and the sample of facilities may suffer from selection issue. Assuming that facilities that were able to implement such management system were also sharing some unobserved characteristics positively corre-

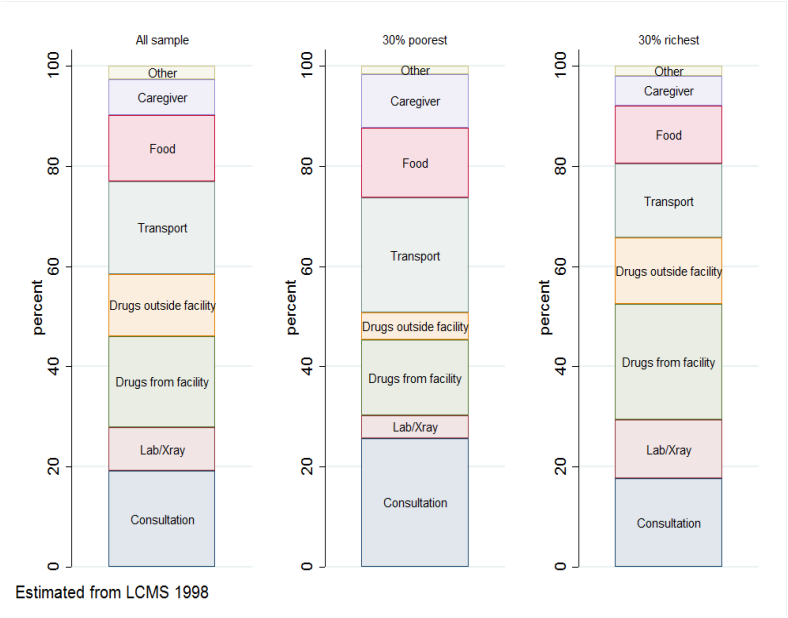
lated with health care use (e.g. better quality of care), the effects on health care utilisation may have been overestimated in those studies. In addition, if the facilities were expecting an alternative funding to compensate for the loss of their revenues and if the funding was a fee for service, then facilities had an incentive to over-report the number of acts provided. Finally, one may want to be cautious of the quality of the causal effect estimated using interrupted time series methods, due to the limited number of data points typically used in these studies, and the relatively poor quality of information deriving from facility registries.

We presented strong contextual elements that may justify our a priori counter-intuitive results. The analysis we presented on the determinants of health-seeking behaviours and OOP medical expenses highlighted that income did not appear to be a main determinant of health care utilisation. We did find that the richest were more likely to seek care but the coefficient associated to income was low. This finding may be attributable to the low level of OOP medical expenses before the policy implementation, where on average, people were spending US$ 2.5 per medical visit, which represents about 1% of their monthly household income. Furthermore, one may want to highlight that the most vulnerable groups (children under-5, elderly, chronically ill and indigents) that are expected to have a higher price-elasticity of demand, were already exempted prior to the implementation of the policy.

Another likely reason for the absence of evidence of effect is that, while user fees were removed for consultation, drugs, laboratory tests and X-rays, these direct medical expenses only represent about half of the total expenses incurred during illness. Detailed information regarding both medical and non-medical health expenses in Zambia was collected in the 1998 LCMS survey, and is presented in Figure 7. This shows that drugs bought outside the facility, transport, food and caregiver costs represent as much as half of the total expenses involved in accessing health care services. In addition, these expenses do not include the opportunity cost of time that has been found to be a main barrier for seeking care in rural settings (Lépine and Le Nestour 2013).

The lack of impact of removing fees on health-seeking behaviours is also supported by other studies. For example, a randomised controlled trial measuring the effects of free care in Ghana (Powell-Jackson et al. 2014), found that while children assigned to the treatment group had a higher number of medical visits, utilisation only increased by 3.7 percentage points and the number of primary health clinic visits by 0.3 points. Furthermore, the literature on the demand for health services underlines that financial affordability of care is only one of several factors explaining why individuals do not use health services when sick. Amongst the factors associated with limited utilisation of care are several pertinent variables, unavailable in our data, such as the preference for health and modern medicine, cultural beliefs and perceptions affecting health-seeking behaviours (Copo, Pisani et al. 1992). Finally, although there is no available measure of quality of care in our data that would allow us to test this, there is ev-

Figure 7: Composition of direct and indirect expenses incurred when using health services



idence suggesting that quality of care in Zambia was deteriorating when fees were removed (Picazo and Zhao 2009), which may also have decreased the demand for health services, despite their increased affordability. However, because of the absence of data on the quality of care we cannot disentangle whether the absence of evidence of effect on utilisation is mostly explained by the price inelasticity of demand or a downward shift in demand that may have resulted from a decrease in the quality of care.

Several studies have identified implementation challenges or other difficulties in the design of the policy that took place in Zambia. First, the policy was decided relatively suddenly by the President, partly for electoral reasons, and there was limited time to plan the reform carefully and communicate the remits of the reform to all levels of the system. As a result, there was still some confusion amongst local health authorities on how to apply the policy change and whether all fees should be abolished or not (Department for International Development 2010). Implementation was therefore probably not random, however results obtained from health facility routine data suggest that it was the 'better' facilities (more organised, more resourced, more supported) that implemented the policy. As such, this reinforces our finding that there was no evidence of effect, even in areas where implementation rate was high. Furthermore, delays in the disbursement of additional resources that were supposed to compensate health facilities led some of them to limit the universality of free care, to avoid running out of cash and being unable to renew drug stocks (Department for International Development 2010, Carasso, Lagarde et al. 2012). Finally, the health financing reform took place against the backdrop of a critical shortage of health workers in certain rural areas (Carasso, Lagarde et al. 2012, McPake, Witter et al. 2013), meaning that in some areas, the populations would have anticipated a very low quality of services provided, discouraging them from using health care services in the first

place, even for free. Similar challenges on the supply-side have limited the effectiveness of the introduction of many other health programmes, such as the social health insurance programme in Mexico (King, Gakidou et al. 2009). In Nepal, lack of information and knowledge from the more disadvantaged population groups contributed to limit the impact of voucher scheme providing financial incentives to women to deliver in health facilities (Powell-Jackson, Morrison et al. 2009). Considering the high degree of publicity given to the policy, it is however unlikely that the degree of knowledge among population was low in the Zambian context.

The identification of medium- or long-term effects was not possible due to the scale-up of the policy to all "peri-urban" areas of the 18 districts included in the donor pool. The remit of this scale-up, which could not be well identified in the household survey data due to the absence of geographic coordinates, prevented the use of the 2010 LCMS wave to assess the longer-term consequences of the policy. However, as highlighted previously, results from health facility routine data finds that utilisation peaked 6 months after the implementation, therefore the timeline we use is probably ideal to capture an increase in utilisation.

In addition to the impossibility of estimating medium and longer run effects, the study presents several limits. Firstly, some of the limitations are inherent to the use of the synthetic control method. Given the non-parametric nature of the synthetic control method, we have run a sensitivity analysis on the covariates to include and the model associated with the smallest RMSPE was selected. More specifically, we estimated a model with no covariate as well as another model that includes province dummies and district population size. While the coefficients obtained were very close, those models lead to a higher RMSPE and higher confidence intervals and hence are not presented in the paper but are available upon request. Regarding the results of the placebo tests, the number of units in the donor pool is similar to the number of Spanish regions used in Abadie et al. (2003) in their seminal paper on the effect of terrorism in the Basque region. Regarding the selection of the units to include in the donor pool, we followed Abadie et al. (2010) that discarded 12 US states from the donor pool because of contamination issues. Overall, there is no guidance on how to select the units from the donor pools but we believe that confidence intervals from the placebo tests are informative. In fact, the confidence intervals we obtain for utilisation CI=[-0.14; 0.12], are reasonably low and close to the CI we obtain with the PSM combined to the DD estimation presented as a robustness check (CI=[-0.19; 0.10] and CI=[-0.15; 0.10] depending on the matching estimator). The effect that the Zambian government was expecting in implementing this policy is unknown but it is likely that the government would have been disappointed if the policy had an effect that is included in our confidence interval given that the upper bound of the CI would have translated in an increase in utilisation from 57% at baseline to 69%. Also, while a better fit of the pre-treatment trend could have been achieved by using a higher number of pre-treatment periods, LCMS data waves collected before 1998 could not be used because the administrative reform of 1991 that led to the creation of 18 rural districts was only reflected in the LCMS data

from 1998. However, we are confident that pooling district level effects as a weighted average of the inverse of the RMSPE may account for this issue and there was evidence that using three pre-treatment periods lead to RMSPEs approaching zero when using a single treated unit.

In addition to this methodological points, it was not possible from the data to estimate the effect of the policy on the likelihood of having catastrophic health expenditures because the items used to estimate the consumption aggregate in the LCMS data changed over time. While this may not be an issue when defining poverty as an interval of the income distribution, this prevents from constructing the share of health expenditures in total expenditures. Finally, while our study focused on intermediary outcomes only (health-seeking-behaviour and OOP medical expenses), the absence of effects on health care utilisation may mean that there was no short-term effect on health outcomes. One may still argue that health outcomes could have been improved if the low level of OOP health expenses among the poor illustrates self-rationing. In which case, the policy could have resulted in a higher consumption of health services by those populations, which could have improved their health outcomes. The low effect of free care on the frequency of medical visits (Powell-Jackson et al, 2014) and the fact that even a positive impact on utilisation of services not necessarily improve health outcomes (Ansah et al. 2009; Powell-Jackson et al., 2014) may however cast doubt on this point.

This study suggests interesting and counter-intuitive lessons for policy-makers, with regards to the equity effects of removing fees. If removing fees does not increase utilisation of care, and in particular that of the poorest groups, but reduces OOP expenses of those already using the services, then the beneficiaries of the policy change are mostly individuals who were already using services. These inequalities associated with use of services imply that removing fees in rural Zambia disproportionally benefited the richer groups, through large and significant reductions in their OOP expenses. In fact, we also showed evidence that given that the richest individuals had higher OOP expenses in absolute terms, the savings generated by the policy was more than three times higher for the richest groups. The conclusion that user fee removal in Zambia was anti-redistributive, although seemingly counter-intuitive, is actually a typical problem of policies promoting universal access to services in settings where initial inequalities are large and barriers to accessing services for the poor multiple (Gwatkin and Ergo 2011).

8 Conclusion

Despite the momentum in policies removing user fees, findings from our study draw attention to the limited benefits that can actually be derived from such a consensual policy decision. The fact that some vulnerable populations benefited from free care before the policy, the fact that the policy might not have reduced all the expenses incurred when seeking care and the existence of non-financial barriers to access to health services are probably the main reasons for

this absence of evidence of effect. It would be interesting to know whether removing these other barriers (e.g. through financial or non-financial incentives) while maintaining user fees would be more effective than removing fees. Given that programmes offering conditional financial or non-financial incentives have also gained a growing attention from policy-makers in the past decade, opportunities are likely to arise that will allow researchers to evaluate their effects, in combination with or in the absence of fee removal.

9 References

Abadie, A. (2005). Semiparametric difference-in-differences estimators. The Review of Economic Studies 72(1): 1-19.

Abadie, A., A. Diamond and J. Hainmueller (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. Journal of the American Statistical Association 105(490).

Akin, J. S., N. Birdsall and D. M. De Ferranti (1987). Financing health services in developing countries: an agenda for reform, World Bank Publications.

Ansah, E. K., S. Narh-Bana, S. Asiamah, V. Dzordzordzi, K. Biantey, K. Dickson, J. O. Gyapong, K. A. Koram, B. M. Greenwood, A. Mills and C. J. M. Whitty (2009). Effect of Removing Direct Payment for Health Care on Utilisation and Health Outcomes in Ghanaian Children: A Randomised Controlled Trial. PLoS Med 6(1).

Blundell, R. and M. C. Dias (2002). Alternative approaches to evaluation in empirical microeconomics. Portuguese Economic Journal 1(2): 91-115.

Burnham, G. M., G. Pariyo, E. Galiwango and F. Wabwire-Mangen (2004). Discontinuation of cost sharing in Uganda. Bulletin of the World Health Organization 82(3): 187-195.

Carasso, B., N. Palmer, M. Lagarde, C. Cheelo, C. Chama, G. Pollen, T. Ngulube, C. Chansa and L. Gilson (2010). Evaluating the impact of abolishing user fees in Zambia- Summary of findings.

Carasso, B. S., M. Lagarde, C. Cheelo, C. Chansa and N. Palmer (2012). Health worker perspectives on user fee removal in Zambia. Human resources for health 10(1): 40.

Cheelo, C., C. Chama, G. Pollen, B. Carasso, N. Palmer, D. Jonsson, M. Lagarde and C. Chansa (2010). Do user fee revenues matter: assessing the influences of the removal of user fees on health financial resources in Zambia, Department of Economics, The University of Zambia.

Cochran, W. G. (1968). The effectiveness of adjustment by subclassification in removing bias in observational studies. Biometrics: 295-313.

Cohen, J. and P. Dupas (2010). Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment. Quarterly Journal of Economics 125(1): 1.

Coppo, P., L. Pisani and A. Keita (1992). Perceived morbidity and health behaviour in a Dogon community. Social science & medicine 34(11): 1227-1235.

Creese, A. L. (1991). User charges for health care: a review of recent experience. Health policy and planning 6(4): 309-319.

Crump, R. K., V. J. Hotz, G. W. Imbens and O. A. Mitnik (2008). Nonparametric tests for treatment effect heterogeneity. The Review of Economics and Statistics 90(3): 389-405.

De Ferranti, D. (1985). Paying for health services in developing countries. World Bank staff working paper (721).Department for International Development (2010). Removing user fees: what matters in making it work? London.

Devadasan, N., B. Criel, W. Van Damme, K. Ranson and P. Van der Stuyft (2007). Indian community health insurance schemes provide partial protection against catastrophic health expenditure. BMC Health Services Research 7(1): 43.

Dube, A. and B. Zipperer (2013). Pooled Synthetic Control Estimates for Recurring Treatments: An Application to Minimum Wage Case Studies. Unpublished paper.

Dzakpasu, S., T. Powell-Jackson and O. M. Campbell (2013). Impact of user fees on maternal health service utilization and related health outcomes: a systematic review. Health policy and planning 29(7):1-14

Fafchamps, M. and B. Minten (2007). Public service provision, user fees and political turmoil. Journal of African Economies 16(3): 485-518.

Gertler, P., L. Locay and W. Sanderson (1987). Are user fees regressive?: The welfare implications of health care financing proposals in Peru. Journal of Econometrics 36(1): 67-88.

Gilson, L. (1997). The lessons of user fee experience in Africa. Health policy and planning 12(3): 273-285.

Gilson, L., S. Russell and K. Buse (1995). "The political economy of user fees with targeting: developing equitable health financing policy." Journal of international Development 7(3): 369-401.

Gobillon, L., Magnac, T. (2014). Regional policy evaluation: Interactive fixed effects and synthetic controls. Mimeo.
Government of the Republic of Zambia (1991). National Health Policies and Strategies.

Gwatkin, D. R. and A. Ergo (2011). "Universal health coverage: friend or foe of health equity?" The Lancet 377(9784): 2160-2161.

Heckman, J. J., H. Ichimura and P. Todd (1998). Matching as an econometric evaluation estimator. The Review of Economic Studies 65(2): 261-294.

Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. Review of Economics and statistics 86(1): 4-29.

King, G., E. Gakidou, K. Imai, J. Lakin, R. T. Moore, C. Nall, N. Ravishankar, M. Vargas, M. M. Tllez-Rojo and J. E. H. vila (2009). Public policy for the poor? A randomised assessment of the Mexican universal health

insurance programme. The lancet 373(9673): 1447-1454.

Kreif, N., R. Grieve, D. Hangartner, A. Turner, J,, S. Nikolova and M. Sutton (2014). The synthetic control method compared to difference-in differences for evaluation health policies. Unpublished manuscript.

Lagarde, M., H. Barroy and N. Palmer (2012). Assessing the effects of removing user fees in Zambia and Niger. Journal of health services research & policy 17(1): 30-36.

Lagarde, M. and N. Palmer (2008). The impact of user fees on health service utilization in low-and middle-income countries: how strong is the evidence? Bulletin of the World Health Organization 86(11): 839-848.

Lavy, V. and J. M. Quigley (1993). Willingness to pay for the quality and intensity of medical care. Low-income households in Ghana. University of California. Berkeley.

Lépine, A. and A. Le Nestour (2013). The Determinants of Health Care Utilisation in Rural Senegal. Journal of African Economies 22(1): 163-186.

Masiye, F., B. M. Chitah, P. Chanda and F. Simeo (2008). Removal of user fees at primary health care facilities in Zambia: a study of the effects on utilisation and quality of care. Regional Network for Equity in Health in east and souther Africa Discussion Paper(57).

Masiye, F., V. Seshamani, C. Cheelo, C. Mphuka and K. Odegaard (2005). Health care financing in Zambia: A study of the possible policy options for implementation, UNZA/MoH.

McPake, B., S. Witter, T. Ensor, S. Fustukian, D. Newlands, T. Martineau and Y. Chirwa (2013). Removing financial barriers to access reproductive, maternal and newborn health services: the challenges and policy implications for human resources for health. Human Resources for Health 11(1): 46.

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

Ministry of Health (2006). Guidelines removal of user fees in public health facilities in Zambia.

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

Onde, R. (2009). Impact of abolishment of user fees in rural health centres: a case of Chongwe district, The Copperbelt University.Picazo, O. F. and F. Zhao (2009). Zambia health sector public expenditure review: Accounting for resources to improve effective service coverage, World Bank Publications.

Powell-Jackson, T., K. Hanson, C. J. M. Whitty and E. K. Ansah (2014). Who benefits from free healthcare? Evidence from a randomized experiment in Ghana. Journal of Development Economics 107(0): 305-319.

Powell-Jackson, T., J. Morrison, S. Tiwari, B. D. Neupane and A. M. Costello (2009). The experiences of districts in implementing a national incentive programme to promote safe delivery in Nepal. BMC health services research 9(1): 97.

Rosenbaum, P. R. and D. B. Rubin (1984). Reducing bias in observational studies using subclassification on the propensity score. Journal of the American Statistical Association 79(387): 516-524.

Sauerborn, R., A. Nougtara and E. Latimer (1994). The elasticity of demand for health care in Burkina Faso: differences across age and income groups. Health Policy and Planning 9(2): 185-192.

Thornton, R. L. (2008). The demand for, and impact of, learning HIV status. The American economic review 98(5): 1829.

Thornton, R. L., L. E. Hatt, E. M. Field, M. Islam, F. Sols Diaz and M. A. Gonzlez (2010). Social security health insurance for the informal sector in Nicaragua: a randomized evaluation. Health economics 19(S1): 181-206.

Wilkinson, D., E. Gouws, M. Sach and S. S. A. Karim (2001). Effect of removing user fees on attendance for curative and preventive primary health care services in rural South Africa.

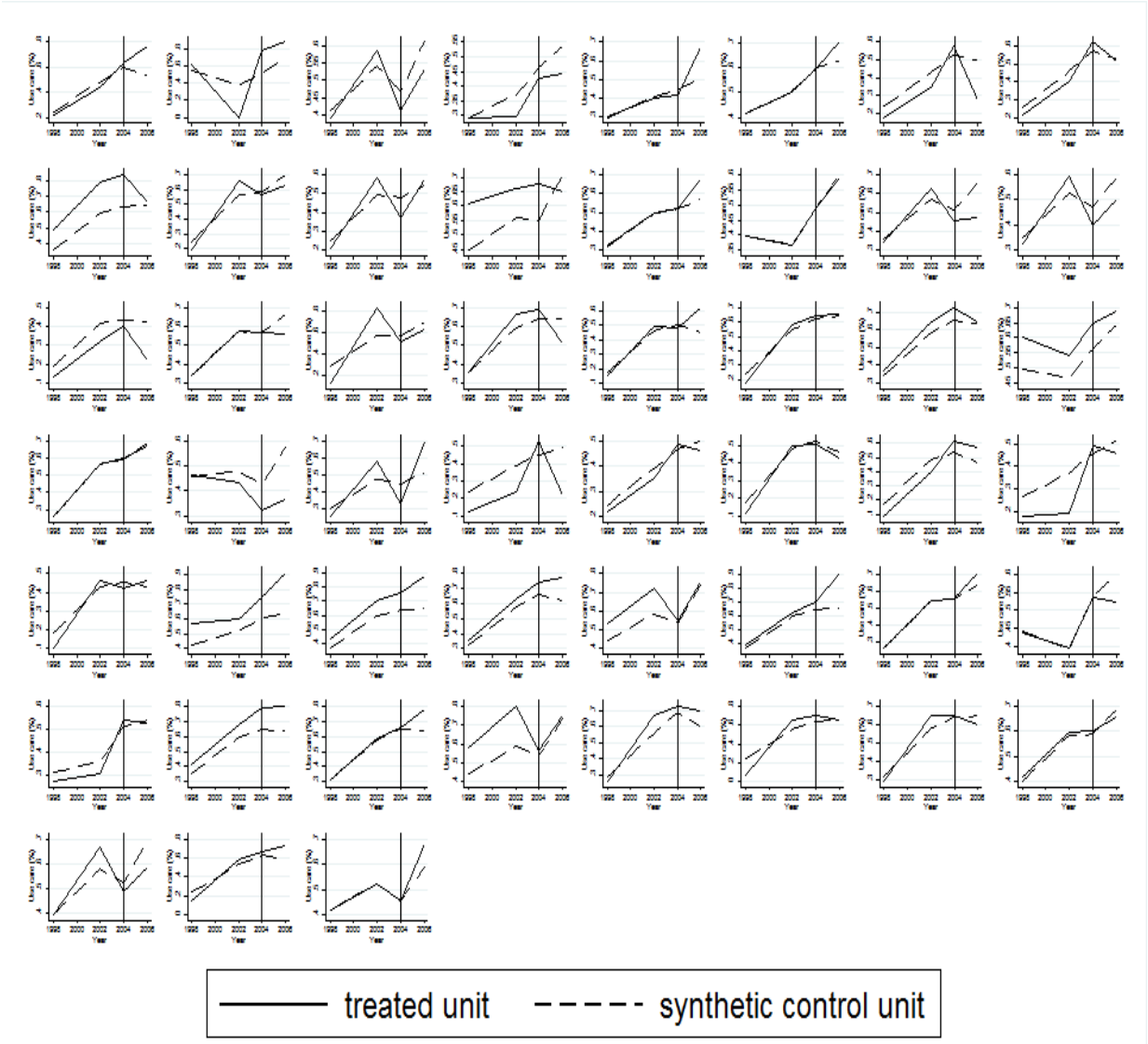
Bulletin of the World Health Organization 79(7): 665-671.World Health Organisation (2010). Health systems financing: the path to universal coverage. World Health report. Geneva, WHO.

Xu, K., D. B. Evans, P. Kadama, J. Nabyonga, P. O. Ogwal, P. Nabukhonzo and A. M. Aguilar (2006). Understanding the impact of eliminating user fees: utilization and catastrophic health expenditures in Uganda. Social Science & Medicine 62(4): 866-876.

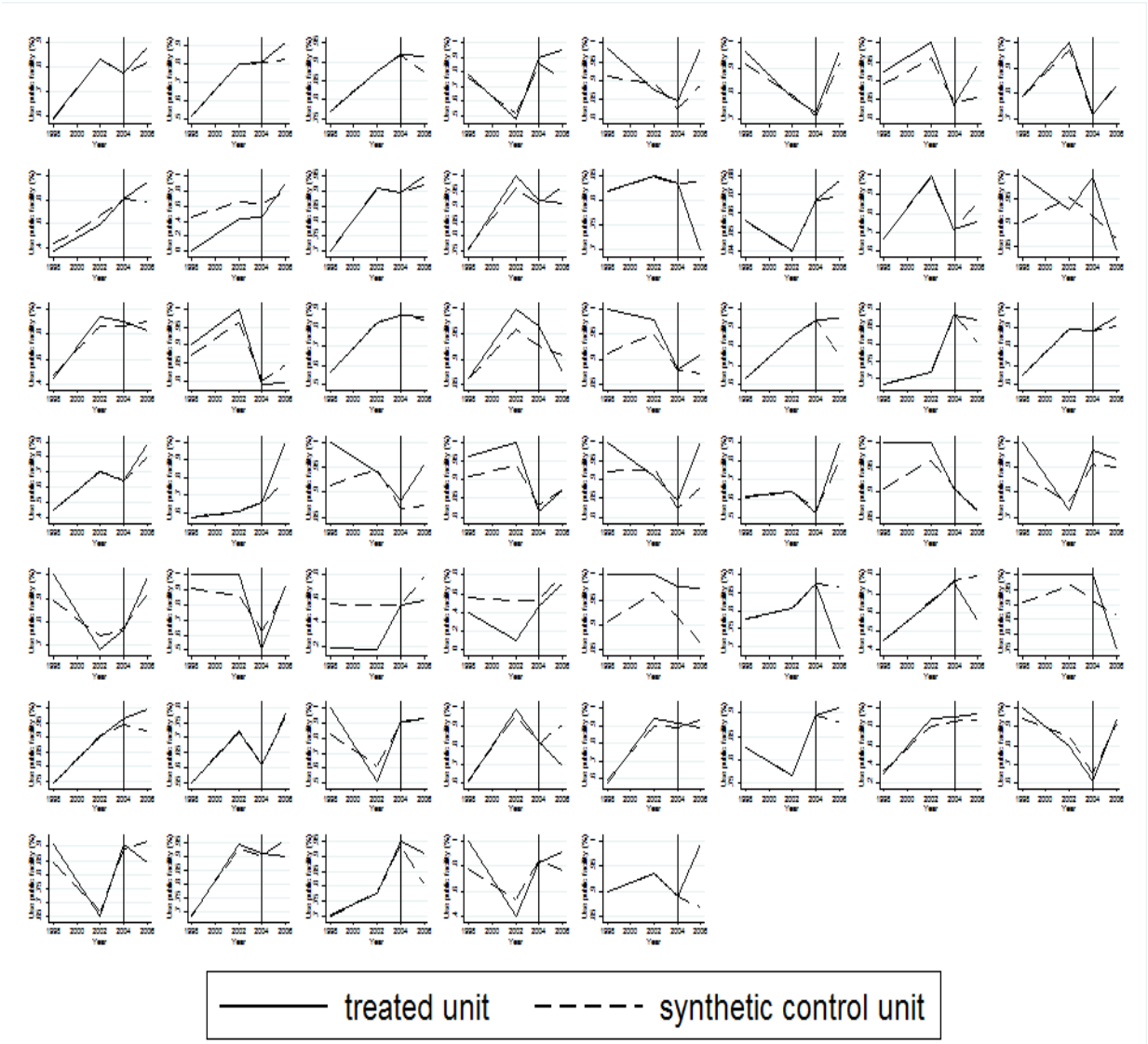
Yates, R. (2009). Universal health care and the removal of user fees. The Lancet 373(9680): 2078-2081.

10 Appendices

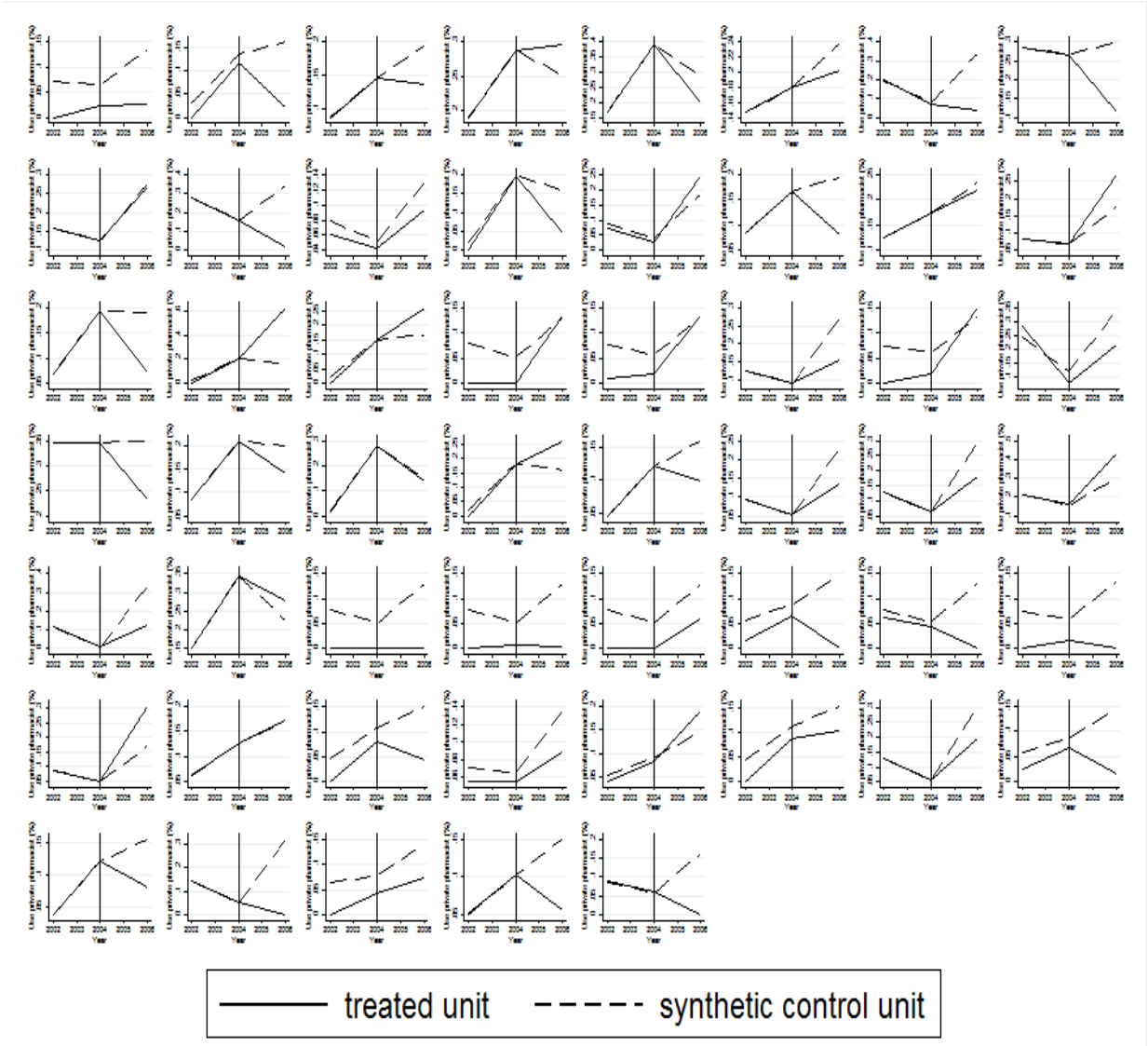
Appendix 1: Trends in treated unit and in its synthetic control for health care use in each treated district



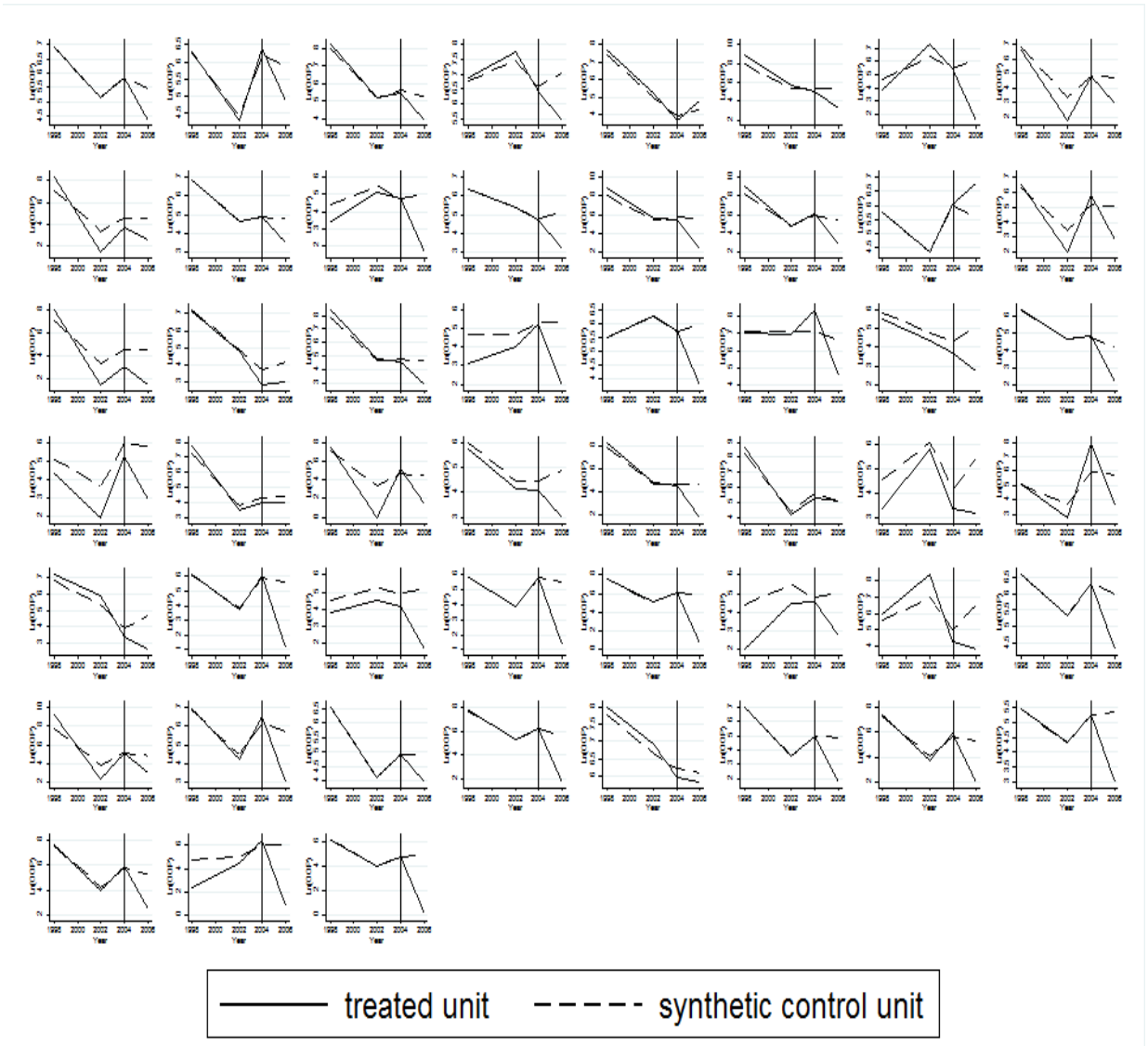
Appendix 2: Trends in treated unit and in its synthetic control for going to a public health facility in each treated district



Appendix 3: Trends in treated unit and in its synthetic control for buying drugs from private sector in each treated district



Appendix 4: Trends in treated unit and in its synthetic control for OOP expenses in each treated district



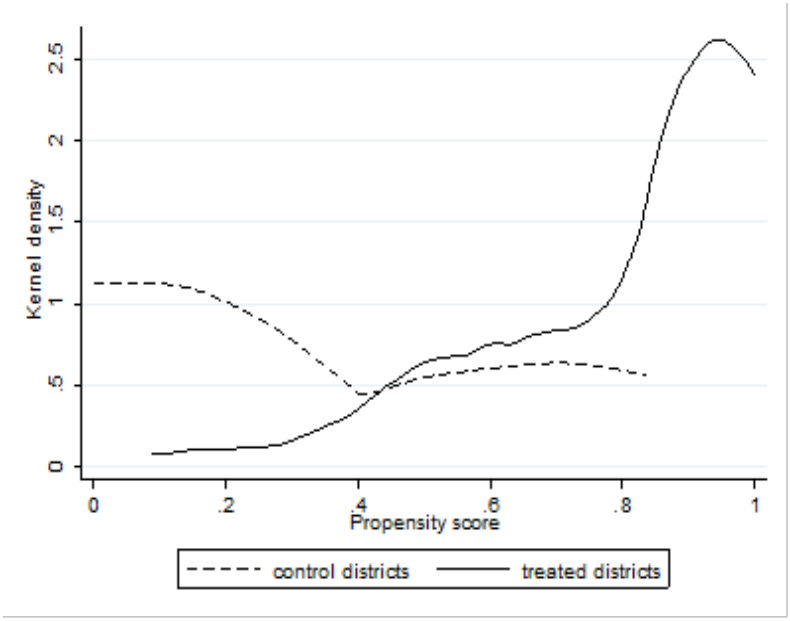
Appendix 5a: Outcome predictor means in the treated unit and in its synthetic control

	Seekcare		Public		Private		Ln(OOP)	
	Treated	Synthetic	Treated	Synthetic	Treated	Synthetic	Treated	Synthetic
Male (%)	0.492	0.487	0.492	0.491	0.492	0.468	0.492	0.492
Female head (%)	0.141	0.115	0.141	0.120	0.105	0.095	0.141	0.118
Rural (%)	0.874	0.612	0.874	0.378	0.863	0.681	0.874	0.551
Median expenditure	13.178	13.474	13.178	13.491	13.145	13.401	13.178	13.372
Median age	20.333	19.826	20.333	20.393	20.500	19.610	20.333	20.284
Median distance to facility	4.333	2.854	4.333	1.660	4.500	2.933	4.333	2.161
Median household size	7.333	8.498	7.333	8.287	8.000	9.038	7.333	8.333
Outcome (1998)	0.310	0.310	0.699	0.699	_	_	6.528	6.523
Outcome (2002)	0.528	0.529	0.827	0.826	0.077	0.077	4.348	4.342
Outcome (2004)	0.547	0.548	0.823	0.822	0.128	0.127	5.196	5.189

Appendix 5b: District weights in the synthetic control

District	Seekcare	Public	Private	Ln(OOP)
Kabwe	0.017	0.059	0.004	0.034
Chililabombwe	0.024	0.041	0.01	0.032
Chingola	0.031	0.029	0.003	0.106
Kalulushi	0.025	0.07	0.126	0.014
Kitwe	0.039	0.051	0.003	0.03
Luanshya	0.024	0.167	0.002	0.026
Mufulira	0.03	0.039	0.006	0.052
Ndola	0.03	0.057	0.004	0.029
Chipata	0.033	0.049	0.014	0.045
Mansa	0.076	0.143	0.017	0.413
Lusaka	0.031	0.048	0.001	0.013
Mbala	0.095	0.04	0.009	0.045
Solwezi	0.076	0.044	0.017	0.06
Choma	0.408	0.077	0.766	0.041
Livingstone	0.017	0.051	0.003	0.016
Mazabuka	0.045	0.034	0.014	0.043

Appendix 6.a: Overlap in densities of the estimated propensity scores for treated and non-treated districts



Appendix 6.b: Propensity score: Effect of district characteristics on the probability of removing user fees (n=72)

	Coef	SE
Rural (%)	4.692*	2.547
Children under-5 with clinic card (%)	-3.469*	2.075
Immunised against DPT (%)	-11.489*	6.101
Median age	0.118	0.293
Median distance to health facility (km)	-3.344	1.277
Median expenditures	-0.001***	0.117
Civil servant (%)	8.953	9.687
Electricity access (%)	6.881	5.234
Access to drinking water (%)	-0.528	1.496
Flush toilet (%)	-6.286	4.735
R-squared	0.55	

*** p<0.01, ** p<0.05, * p<0.1. Probit estimation. The dependent variable is a dummy equal to one if the district is designated to remove user fees and zero otherwise.