UNU-WIDER | 30 YEARS OF RESEARCH FOR DEVELOPMENT

WIDER Working Paper 2015/108

General budget support, health expenditures, and neonatal mortality rate

A synthetic control approach

Kaisa Alavuotunki*

November 2015

United Nations University World Institute for Development Economics Research

wider.unu.edu

Abstract: I examine impacts of general budget support in 12 countries using the synthetic control approach. First, I analyse changes in government expenditures on health before and after the introduction of budget support. Second, I look at neonatal mortality (a presumed proxy for improvements in health services), concentrating on the countries that increased their health spending following the introduction of general budget support. The results indicate that, at least in Tanzania, Burkina Faso, Rwanda, and Malawi, budget support did have a positive effect on health spending and, except for Rwanda, neonatal mortality rate declined relatively faster than in the synthetic control countries.

Keywords: general budget support, health sector expenditures, neonatal mortality, developing countries, synthetic control method
JEL classification: H51, F35, O57

Acknowledgements: I would like to thank Pertti Haaparanta, Jukka Pirttilä, and Saurabh Singhal as well as seminar audiences at the UNU-WIDER 30th Anniversary Conference 2015, the Nordic Conference on Development Economics 2015, and the Finnish Economic Association XXXVII Annual Meeting 2015 for their useful comments. I am grateful for support from the Yrjö Jahnsson Foundation and Academy of Finland.

Tables and Figures at the end of the paper.

*Aalto University, Helsinki, Finland; kaisa.alavuotunki@aalto.fi.

This study has been prepared within the UNU-WIDER project 'The Economics and Politics of Taxation and Social Protection'.

ISSN 1798-7237 ISBN 978-92-9230-997-8 https://doi.org/10.35188/UNU-WIDER/2015/997-8

Typescript prepared by Judy Hartley for UNU-WIDER.

UNU-WIDER gratefully acknowledges the financial contributions to the research programme from the governments of Denmark, Finland, Sweden, and the United Kingdom.

1 Introduction

General budget support (GBS) started gaining popularity in the early 2000s as a reaction to prevailing problems of project aid and other existing aid modalities. The problems concerned both the earlier versions of programme aid (balance of payments and import support) and the major aid modality at the time, project aid. While programme aid was criticized for being ineffective, project aid faced challenges with local ownership, harmonization, and predictability (Collier et al. 1997; Disjkstra 2013; Dollar and Svensson 2000; IDD and Associates 2006). GBS was to be used in a selected group of countries (most of which were also highly indebted poor countries) channelling aid directly to recipient country government budgets using their own allocation, procurement and accounting systems, without being linked to specific project activities.

One of the desired goals of general budget support is to enable recipient country governments to allocate more resources into pro-poor sectors, such as health, education and agriculture. Once received, GBS becomes a part of recipient country's budgetary processes. The funds are allocated based on the recipient government's principles and, despite the inbuilt requirement for donor–recipient dialogue, donor countries have little control over the use of those funds (Dijkstra 2013; Dijkstra and de Kemp 2015; Molenaers et al. 2010). The non-earmarked nature of GBS makes it fairly difficult to evaluate whether or not these additional funds have been allocated in a pro-poor manner or to distinguish the impact of GBS from other government revenue sources. According to existing cross-country studies, if anything, receiving GBS has had a positive impact on government health expenditures (Alavuotunki 2015; Dijkstra and de Kemp 2015; WHO 2010a). Needless to say, among GBS-receiving countries, the spectrum of different responses is wide and it is practically impossible to reveal the impact of GBS on an individual country with a cross-country estimation framework.

Thus, the main contribution of this paper is to gain deeper understanding of the heterogeneous effects of GBS on a country-specific level. I use the synthetic control method to explore whether getting non-earmarked GBS funds has had an impact on the size of health expenditures in an individual country framework. Further, just looking at the budget allocation would not reveal much about whether additional funds were actually used to improve services for people. Thus, I select neonatal mortality rate as a health outcome indicator that is used as a presumed proxy for the improved health services (both access and quality). Since very simple interventions in improving access to health services and quality of the services (i.e. more skilled workers, better access) can help reduce neonatal mortality, accelerated reduction of the neonatal mortality rate can be seen as an early indicator of health policy or programme success (Bhutta et al., 2010; Lawn et al. 2008; Rajaratnam et al. 2010).

While the synthetic control method helps observe whether level of health sector spending or declining rate of neonatal mortality are any different in a GBS-recipient country after the introduction of GBS, the method does not statistically examine the connection between increased health sector spending and decreased neonatal mortality; rather it looks at both as a separate outcome of GBS. I am, nevertheless, interested to see if these two have happened hand in hand. A result where government health sector spending increases without any significant impact on neonatal mortality reduction rate should, at the very least, indicate a failure in the health policy design.

The synthetic control method was created by Abadie et al. (Abadie and Gardeazapal 2003; Abadie et al. 2010, 2014) as an attempt to combine quantitative and qualitative methods for the use of empirical political research. For an individual country, assessing impacts of different

interventions is often challenging due to the lack of suitable control countries. In other words, defining a counterfactual (what would have happened for the observed variable if the intervention did not take place) is problematic. The synthetic control method offers a systematic and transparent way to choose comparison units for a case country. The method is based on the idea that a combination of control units (a so-called 'synthetic control') corresponds better to the observed country than any single 'real' country. The benefits of using the method over, for example, running a regression analysis, are avoidance of extrapolation and the opportunity to examine explicitly each individual counterfactual. The latter, especially, helps create deeper understanding of country groups and forms the basis for the use of qualitative analysis together with quantitative (Abadie et al. 2014).

Twelve of the top GBS-receiving countries in the 2000s are selected for the analysis.[1] The results indicate that in Burkina Faso, Rwanda, Tanzania, Malawi and, to some extent, in Niger and Burundi, the expansion of health sector financing has been indisputable after receipt of GBS, while in Mozambique, Ghana, and Uganda there has been some positive development but it is unclear whether that can be attributed to GBS. I also show that the countries with increased health sector spending, especially Burkina Faso, Malawi, Tanzania, and Burundi, have also managed to lower their neonatal mortality rate more than their synthetic controls.

The rest of the paper is organized as follows. Section 2 provides an overview of the literature on the synthetic control method as well as the relevant research on GBS and government budget allocation. Section 3 introduces the methodology behind constructing a synthetic control. Section 4 presents the data and Section 5 reports the results. Section 6 concludes.

2 Literature review

Until recently the academic literature on the effects of GBS has been quite scarce.[2] Dijkstra and de Kemp (2015) provide a good overview of different attempts to evaluate GBS. They list four major challenges in assessing the results of budget support: (1) identifying a counterfactual; (2) establishing attribution; (3) the variety of objectives; and finally (4) diverse donor priorities, against which the success of budget support is judged.

On the impacts of GBS, a few cross-country econometric studies exist, mainly on the effects of GBS on budget allocation (Alavuotunki 2015; Antunes et al. 2013; IOB 2012; WHO 2010a), growth (Alavuotunki and Sandström 2015; Bigsten et al. 2011) and human development index (Beynon and Dusu 2010). Previous academic work mostly focused on the average effect of receiving GBS on government budget allocations, failing to take account of the possible heterogeneity of the effect in different recipient countries.

The results of the limited literature on the impacts of GBS on health sector spending are mixed. While IOB (2012) report a positive and statistically significant impact of GBS on both health expenditure share of total government spending and on health sector spending per GDP, Antunes et al. (2013) conclude that, even though an increase in total government spending has a positive effect on health sector spending, GBS funding has no observed direct impact on government health sector spending other than one that comes through an increase of total

[1] Mozambique, Tanzania, Rwanda, Sierra Leone, Malawi, Burundi, Uganda, Madagascar, Niger, Burkina Faso, Ghana and Mali (Table 1).

[2] A search for 'general budget support' in the title, abstract or keywords in one of the top journals in the field, *Journal of Development Economics*, found one hit.

government spending. In my previous research (Alavuotunki 2015), I find that while GBS is associated with higher health sector spending, it has more to do with donor GBS allocation criteria than the help of additional funding. WHO (2010a) report a small but significant effect of GBS/capita on the total health expenditures of a government based on a six-year panel of 79 countries.

Łukasz (2014) studies causal links between on-budget aid (includes all on-budget aid, not only GBS) and government expenditures using Granger causality in heterogeneous panels. His results suggest that aid (in general) substitutes for government revenue and that recipient governments do not actually increase spending (no quantity effect) with increased aid but rather change the way they use their own resources. The general effects of aid on government budget allocation have also been studied by Fosu (2010) and Fosu and Quattri (2012), indicating that aid has a positive effect on public investments but their results do not say anything specific about GBS.

Several country-specific non-academic evaluations analyse the impact and outcomes of GBS. Joint evaluations by the European Commission and Associates have covered countries such as Tanzania, Mozambique, Mali, South Africa, and Tunisia (Caputo et al. 2011; Caputo et al. 2013; Lawson et al. 2007; Lawson et al. 2011; Lawson et al. 2013; Lawson et al. 2014) and evaluation financed by the Ministry of Foreign Affairs of the Netherlands extensively studied GBS in Ghana, Mali, Nicaragua, Tanzania, Vietnam, and Zambia (De Kemp et al. 2011; Dijkstra 2013; IOB 2012). In general, these country-specific evaluations suggest that general and sector budget support have had a positive influence on public spending by increasing funds channelled to priority sectors,[3] mainly to education and health sectors (IOB 2012; Lawson et al. 2013; Lawson et al. 2014). For example, in Tanzania, spending on its six priority sectors (health, education, water, agriculture, energy, roads) doubled during the evaluation period (from 2005 to 2011) and the priority sector share of all public spending increased from 40 per cent to 50 per cent of total spending (Lawson et al. 2013). The evaluations also conclude that GBS had an effect on non-income poverty, especially on enrolment rates to primary and secondary schools but income poverty reductions cannot be attributed to GBS funds. The challenge of a case study approach is the lack of a rigorous methodology and a tendency to overlook the endogeneity issues when following a logical framework from inputs to outputs and outcomes (Dijkstra and de Kemp 2015; Elbers and de Hoop 2009).

Most of the current literature on the impacts of GBS suffers from one of two major weaknesses. First, in their attempt to describe associations between GBS and government expenditures, country-specific studies fail to say anything about the counterfactual, i.e. what would have happened in a country if the intervention (receiving GBS) had not taken place. Second, while well-conducted cross-country studies may solve some of the endogeneity and reverse causality issues through instrumentation strategies (Dijkstra and de Kemp 2015), they still, at their best, reveal only an average effect of GBS amongst the recipient countries. In the case of heterogeneity, cross-country estimations are not able to reveal much about the effect in a single country. The synthetic control method used in this paper can possibly overcome both of the weaknesses by allowing us to analyse one specific country at a time and also to construct a synthetic control that helps to identify a counterfactual.

[3] Priority sectors are defined by each recipient government.

3 Methodology

3.1 Synthetic control method

The synthetic control method was first used by Abadie and Gardeazabal (2003) in evaluating economic development in the Basque Country in the absence of the terrorism. The method was further developed in Abadie, Diamond and Hainmueller (2010—ADH 2010 from now on) and Abadie, Diamond and Hainmueller (2015) and applied in several other comparative case studies (Billmeier and Nannicini 2013; Kirkpatrick and Bennear 2014; Mideksa 2013; Singhal and Nilakantan 2012). Abadie and Gardeazabal (2003) reveal that the terrorist conflicts in the Basque Country caused 10 percentage points lower GDP per capita in the Basque Country relative to a synthetic control region without terrorism. In ADH 2010, the authors test the applicability of the synthetic control method in a comparative case study and take a closer look at the effects of California's Tobacco Control Programme on cigarette sales, the results suggesting that the effects of the tobacco control programme are much larger than prior estimates had reported. In Abadie et al. (2014), the use of synthetic controls in small sample comparative studies is illustrated by studying the economic impact of German unification on West Germany's economy.

The synthetic control method has served researchers in a quite diverse set of topics. Billmeier and Nannicini (2013) apply the method to assess the effect of trade liberalization on economic growth in developing country contexts. They conclude that liberalization has a positive effect on the economy but the effect is less positive especially if the liberalization took place in the 1990s and in Africa. Mideksa (2013) applies the synthetic control method to explore the economic impact of natural resource endowment on the Norwegian economy and reports that about 20 per cent of the annual GDP per capita increase is due to the endowment of petroleum resources. Kirkpatrick and Bennear (2014) use the synthetic control together with the difference-in-difference method to evaluate the effect of property-secured loans to homeowners on clean energy investment. Nilakantan and Singhal (2012), in turn, apply the method to evaluate an effect of a specially trained police force dedicated to combat violence in one of the Indian states.

There are a few attempts to use synthetic controls to examine the development of specific issues; however, these are yet to be published. Among these attempts are Gathani, Santini and Stoelinga (2013) who use the method to examine the impact of a so-called 'one-stop shop' in Rwanda on new company registrations, and Lépine et al. (2015) who estimate the short-term effects of user fee removal in primary care in Zambia by using a pooled synthetic control method.

The synthetic control approach reproduces an outcome trend for the country of interest by convex combination of control units and uses it to construct a counterfactual (what would have happened in the absence of intervention). In theory, the created synthetic control unit should have identical outcome behaviour in the pre-treatment period with the treated country, thus implying that the post-treatment difference in the behaviour of the outcome between the treatment and synthetic control countries can be explained by the treatment.

Let X_1 be a vector of pre-treatment outcome and control variables for the GBS-receiving treated country and X_0 be the corresponding matrix for the J possible control countries. Then a synthetic control weight matrix W is chosen to minimize $(X_1 - X_0W)'V(X_1 - X_0W)$, where V is a diagonal matrix reflecting the relative importance of the different X's. Weight matrix W chooses countries contributing to the synthetic control unit. Matrix V ensures that the chosen synthetic unit matches the treated unit in its pre-intervention outcomes (minimizing the mean

squared errors). Since *W* is a function of *V* it creates a two-stage simultaneous optimization problem.

In the following, I give a brief overview of the ADH (2010) methodology using the same notation and equations as used by the authors.

Let *J+1* be the number of countries in our pool out of which *only one* is exposed to a treatment at time T_0 and $1 \leq T_0 < T$. In our case, a country starts receiving GBS at T_0 and continues receiving at least until T, the end of the observed period. Y_{it}^N is the observed outcome for each country *i* at the time *t* when not exposed to the treatment. Y_{it}^I is the observed outcome when the unit is exposed to the intervention. The intervention should not have any effect prior to its implementation.

The effect of the treatment on outcome is defined by $\propto_{it} = Y_{it}^I - Y_{it}^N$ for a country unit *i* at time any *t* and D_{it} is a dummy variable that takes a value of one in time *t* of the treatment and zero otherwise. Thus the observed outcome for country *i* at time *t* is defined by:

$$Y_{it} = Y_{it}^N + \propto_{it} D_{it} .$$

Only one country exposed to the intervention is analysed at a time thus:

$$D_{it} = \begin{cases} 1, & i = 1 \; and \; t > T_0, \\ 0, & otherwise. \end{cases}$$

Since outcome variable Y_{1t}^N is not observable when $i = 1 \text{ and } t > T_0$, it needs to be estimated in order to be able to calculate treatment effects $\propto_{1t}$ for $t > T_0$. Solving Y_{1t}^N means solving a counterfactual for the treatment, i.e. 'what would have happened in the absence of the treatment?'.

Now, let Y_{1t}^N be defined by a generalized difference-in-difference model with fixed effects where country-specific effects are allowed to vary in time:[4]

$$Y_{1t}^N = \delta_t + \theta_t Z_i + \lambda_t \mu_i + \varepsilon_{it}$$

where δ_t is a time-varying unknown common factor that is constant for all the units across time and Z_i is the vector of all observed covariates. Covariates should not be affected by the intervention but they may be also time-varying, θ_t is a vector of unknown parameters, λ_t is a vector of unobserved common factors and μ_i is a vector of unknown country-specific unobserved confounders. Error terms ε_{it} are unobserved transitory shocks at the country level with zero means.

Now, let vector of weight be $W = (w_2 + \ldots + w_{J+1})'$ where $w_j \geq 0$ and $\sum_2^{J+1} w_j = 1$. Each value of vector *W* presents a potential synthetic control, in other words a weighted average of control countries. For a given *W* the outcome for a synthetic control is:

$$\sum_2^{J+1} w_j Y_{jt} = \delta_t + \theta_t \sum_2^{J+1} w_j Z_j + \lambda_t \sum_2^{J+1} w_j \mu_j + \sum_2^{J+1} w_j \varepsilon_{jt} \qquad (1)$$

[4] In the difference-in-difference model country specifics are time-invariant.

Let there be weights $W^* = (w_2^* + \ldots + w_{J+1}^*)'$ such that:

$$Y_{1t} = \sum_2^{J+1} w_j^* Y_{jt}, \forall t \in = 1, \ldots, T_0 \text{ and } Z_1 = \sum_2^{J+1} w_j^* Z_j \quad (2)$$

The first part of equation (2) shows that the weighted average of the pre-treatment outcome of the control countries equals the pre-treatment outcome of the treated country, while the second part shows that the weighted average of the country-specific pre-treatment characteristics of the control countries replicates the pre-treatment characteristics of the treated country.

Further, ADH (2010) prove that if $\lambda_t' \lambda_t$ is non-singular, then:

$$Y_{1t}^N - \sum_{j=2}^{J+1} w_j^* Y_{jt} = \sum_{j=2}^{J+1} w_j^* \sum_{s=1}^{T_0} \lambda_t \left(\sum_{n=1}^{T_0} \lambda_t' \lambda_t \right)^{-1} + \lambda_s' (\varepsilon_{js} - \varepsilon_{1s}) - \sum_{j=2}^{J+1} w_j^* (\varepsilon_{jt} - \varepsilon_{1t})$$

(3)

They also show that under a set of standard assumptions,[5] the mean of the right-hand side of the equation (3) above is zero if 'the number of pre-intervention periods is large relative to the scale of the transitory shocks'. Thus the estimate of the treatment effect α_{1t} is:

$$\hat{\alpha}_{1t} = Y_{1t}^N - \sum_{j=2}^{J+1} w_j^* Y_{jt} \quad \forall t \in \{T_0 + 1, \ldots, J + 1\}$$

In real life, equation (2) holds only approximately and sometimes it may not hold at all. The latter happens for example when $(Y_{11}, \ldots, Y_{1T_0}, Z'_1)$ does not belong to the convex hull of $\{(Y_{21}, \ldots, Y_{2T_0}, Z'_2), \ldots, (Y_{(J+1)1}, \ldots, Y_{(J+1)T_0}, Z'_{J+1})\}$ and especially if it falls far from it. Later in the analysis, I found that this is, in fact, the case with just some of the treatment countries (for example Mozambique, Mali and Sierra Leone): their neonatal mortality rates rise far above the neonatal mortality rate of any of the control group countries and thus it is technically impossible to calculate the treatment effect without extrapolation.

The synthetic control method allows us to calculate these discrepancies for each set of treatment units and control groups separately and thus judge each set separately against whether the characteristics of the treated unit are sufficiently matched or not. This requirement can be called 'a need for a common support' between the treated and comparison countries and it simply means that there is a need to select a control group where the characteristics of the control group countries are as similar as possible to the treatment country characteristics to avoid large interpolation biases (Billmeier and Nannicini 2013).

To implement the synthetic control method, let *W* be a vector for positive weights that sum to one. Let X_1 be a vector that describes the pre-intervention characteristics (outcome variable and country covariates) of the treated country. Let X_0 be a matrix of the same characteristics for all the countries in the control group that has not been treated. A vector of weights W^* is used to minimize a distance $\|X_1 - W^* X_0\|$. In practice, ADH (2010) show that for the computation of the weights, we need to consider only a few linear combinations of pre-intervention outcomes and check if equation (2) holds approximately for theses weights. The difference between the treated and the synthetic control is $\|X_1 - W^* X_0\| v = \sqrt{(X_1 - W^* X_0)' V (X_1 - W^* X_0)}$, where

[5] ADH (2010) Appendix B.

V is a diagonal matrix reflecting the relative importance of the different Xs. The choice of V influences the weights W and thus the optimal choice of V assigns weights to linear combinations of Z_0 and Z_1 so that the mean squared error of the estimate for the synthetic control minimizes. The choice of V can be data-driven or it can sometimes be based on subjective assessments.[6]

The benefits of the synthetic control method compared to the widely applied difference-in-difference method come, for example, from avoiding extrapolation and allowing time-invariant covariates. In addition, the method is transparent in the sense that all the countries and weights used in forming the synthetic method are laid out in the open and can be given qualitative interpretations. Further, in a practical application, country covariates enter the equations as mean values for the pre-treatment period, allowing flexibility in the development country settings where perfect time series are rarely available. Only outcome variables need to cover each year.

3.2 Limitations of the methodology

One of the main limitations of the synthetic control method is that the standard tests for inference cannot be used. However, there are other methods available. I follow the example of ADH (2010) and Billmeier and Nannicini (2013) and run placebo tests to assess the robustness of the results (elaborated further in Section 3.3). Another disadvantage with the method is that although the synthetic control method can handle endogeneity due to omitted bias, it still suffers from reverse causation, which is the case, for example, if a decision to give GBS was based on the expectation of future growth prospects in health expenditures. This might be the case to some extent in our setting but it cannot be eliminated. Another two assumptions that are necessary if the synthetic control method is to produce unbiased results are the non-interference assumption (i.e. covariates of vector Z_i should not be affected by the intervention) and the no-anticipation effect assumption (the pre-treatment period outcome is not affected by the treatment).

In addition, one has to assume that no other big reforms influencing our outcome variable take place at the same time. Naturally, for example, national health system reform taking place would potentially have a huge impact on the size of health expenditures. It is likely the case that the treatment might include heterogeneous reforms at the country level but these reforms can also be thought to be a consequence of general budget support funding.

The need for a common base may also create problems since the country group I am interested in (countries receiving GBS) is exceptionally poor and, since it is not possible to have other treated countries in the control country pool, the GDP per capita of the countries left in the control pool might not always match well with our treated countries. Looking simply at health spending per capita would lead us to a situation where the outcome variables of our case countries would mostly lie outside the convex combinations of the outcome variables of control pool countries and using the synthetic control approach would not be feasible (violation of equation 2). Thus, health expenditure per GDP is used instead, which allows us to proceed with the analysis.

Another important feature not addressed by the methodology is discussed in Billmeier and Nannicini (2013). Economic reforms rarely happen overnight and, in our case, the impact that GBS might have on government expenditures and, especially, on actual health indicators, might

[6] In this paper, I am using STATA software SYNTH by ADH to perform calculations needed for each case study.

take some time to occur. This kind of measurement error introduces a simple attenuation bias in the results, 'as the effect detected by the SCM would be lower if reforms were diluted across multiple years' (Billmeier and Nannicini 2013: 991).[7]

3.3 Inferential techniques

The synthetic control method does not allow the use of standard (large-sample) inferential techniques due to the nature of comparative case study analysis, with a small number of observations in the control pool and a short time period covered by the sample. However, there are other methods of validation. I follow the example of ADH (2010) and Billmeier and Nannicini (2013) and run the following analysis for each of the potential control countries in every case.

First, the intervention is reassigned systematically to all control units not directly exposed to the intervention (with the actual treatment country in the pool of possible controls), after which these 'placebo' studies are compared to our original results with the actual GBS country as our treated unit. The proportion of estimated placebo effects that are greater than or equal to the one estimated for the unit representing the case of interest is calculated (*p*-value). In the absence of randomization, the *p*-value has an interpretation as the probability of getting an estimate at least as large as the one obtained for the unit representing the case of interest when the intervention is reassigned at random in the dataset (ADH 2010). If the estimated effect for a GBS-receiving country is unusually large relative to the distribution of placebo effects (Abadie et al. 2014), it indicates the effect of GBS is significant.

Another possible placebo test, in time placebo, is not applied here since the pre-treatment time period is relatively short.

4 Data

GBS enters the analysis only to determine a treatment year (when GBS is introduced in the country).[8] All the observed GBS countries have received GBS for the whole treatment period (from the treatment year onwards) and countries that have received GBS at any point in time, even if for only a year, are excluded from the possible control country sample.

Using the synthetic control method requires also that data for the outcome variable have to cover all units and years included in the analysis. The period of analysis is rather short because complete annual data on the outcome variables of interest (government expenditures and neonatal mortality rate) are not generally available before the 1990s. Fortunately, the requirements for covariates are more flexible since the covariates enter the calculations as averages. In other words, all the non-GBS-receiving countries with non-missing time series data on health expenditure (or later in neonatal mortality rates) and a minimum of one value for each covariate can be used as controls. All data sources used in the analysis are described in Appendix Table A1.

---

[7] SCM = synthetic control method.

[8] It might be sometimes quite difficult to determine the exact start year for GBS since it went under different names in the early 2000s. I have, however, assigned each country a start year based on careful study of OECD-CRS datasets and different individual country reports.

4.1 Treatment countries

Appendix Table A1 lists the GBS start year (if any) for each country. The start of GBS as an instrument might in some cases be unclear and is thus categorized into three different periods: 'early 2000s' if GBS started in 2003 or earlier, 'mid-2000s' if GBS started between 2004 and 2006, and 'late 2000s' if GBS started after 2007. To choose the set of treatment countries for the analysis, it makes sense to focus attention on the top GBS-receiving countries following the assumption that the more GBS a country receives, the more likely it is to have an effect on government expenditures or, finally, on the quality of and access to health services (see intervention logic in Figure 1). Table 1 illustrates countries according to the amounts of GBS received in 2005 (GBS/total government expenditures; GBS/total aid; GBS/GDP; GBS in USD). The top 12 countries with the highest GBS per total government expenditure and high GBS per GDP shares are chosen as our case study. These are Rwanda, Sierra Leone, Mozambique, Malawi, Madagascar, Tanzania, Burundi, Uganda, Ghana, Niger, Mali, Burkina Faso. These countries all started receiving GBS in the first half of the 2000s and continued receiving high amounts of it for the whole 'post-treatment' period. Each country's starting year is adjusted individually. The GBS level of the selected countries remains towards the top throughout the whole treatment period.

Further, to provide a pool of potential control countries with enough of a 'common base', it is also necessary that a sufficient set of countries exists in the same economic region. For each case study analysis, only one GBS-receiving country at a time is included while the rest are dropped and not used in the particular analysis. I end up with 55 potential control countries, of which 15 are in sub-Saharan Africa (SSA) and 10 in Asia (South, East or Pacific), 16 in Latin America, 10 in the Middle East and North Africa and 11 in Europe and Central Africa. Of these, countries with full time series data on (1) health expenditure for 1995–2011, (2) the neonatal mortality rate for 1990–2011, and (3) at least one pre-treatment value for each covariate are included in the analysis as potential controls.

4.2 Health expenditure and its covariates

Following the logical framework of expected outcomes and the impact of general budget support funds (Figure 1), one expected impact channel of GBS on the end outcomes (income and non-income poverty reduction) is through increased funding on priority sectors. The health sector is chosen to represent priority sectors due to data availability. Ideally, looking at government health sector spending per capita would seem like the best way to measure how much money is used on healthcare per person. However, since many of the countries receiving GBS are amongst the poorest of the poor, I would end up in a situation where equation (2) does not hold, thus using health expenditure per GDP takes into consideration the different levels of income in the country and allows the use of synthetic controls in determining the counterfactual for countries.

I use panel data covering most of the aid-receiving developing countries and years 1995–2011. Government health expenditures per capita have increased through the years in all regions (Figure 2) but when measured relative to GDP there is more variation in the trend between regions (Figure 3). Figure 4 shows that countries receiving general budget support have on average increased their health expenditures faster than countries not receiving it.

A set of covariates used in the literature (Alavuotunki 2015; Antunes et al. 2013; IOB 2012) of cross-country regressions are chosen, including GDP per capita, age dependency rate (the number of people over 65 and under 16 for a population of persons aged 16–65), the log of population, and government military expenditure per GDP (a proxy for total government

expenditure per GDP). The last one is needed since the level of total government expenditures is an important determinant for the level of health sector expenditure; however, it cannot be included as a covariant since it is directly influenced by GBS funds. On the other hand, government military expenditures are highly correlated with government total expenditure but, according to previous studies, there is no evidence that GBS flows would increase military expenditures (see for example IOB 2012). Other aid flows allocated to the health sector also are possibly important determinants of the level of government health sector spending (governments might be less/more willing to invest in health sector reforms if there are large externally funded projects in place). Yet these flows are likely to be influenced by receiving GBS as some of the donors might start channelling their aid from health sector projects to GBS. Thus this control is left out of the analysis.

4.3 Neonatal mortality rate and its covariates

When assessing whether receiving GBS is followed by improvements in health conditions or, to begin with, improved access to the healthcare system, I need an indicator that is likely to react immediately to improved access to and quality of the existing system. A health indicator that is seen to be fairly sensitive to changes in access to and quality of the healthcare system is the neonatal mortality rate (the number of newborn babies dying before reaching 28 days of age per 1,000 births). One third of all child deaths occur within the first month of life, but providing skilled care to mothers during pregnancy, as well as during and after birth, can greatly contribute to child survival (UNICEF 2012, 2013). One fundamental cause of high neonatal mortality rates is poor access of mothers and newborns to basic health services and the single most important factor in the decline of maternal and newborn deaths worldwide is access to quality care for mothers and newborns (UNICEF 2009, 2011, 2013). Decline in the mortality rate could signal better access to a healthcare system or improvements in the skills of the health sector workers. Thus neonatal mortality rate is chosen as a proxy for improvements in a country's healthcare systems.

For the neonatal mortality rate, data are available for the years 1990–2011. Globally, the neonatal mortality rate has a declining trend (Figures 5 and 6) and, especially in the 2000s, the vast majority of the developing world (with a few exceptions such as Lesotho, Zimbabwe, and Botswana) has managed to decrease the number of newborn deaths. Figure 7 plots trends in the neonatal mortality rate in GBS-receiving countries, in the rest of the developing world and in African countries not receiving GBS.

In the case of the neonatal mortality rate, covariates include GDP per capita level, proportion of the total population aged under 15, adolescent fertility rate and access of female students to some education.[9] It is also very likely that there are many other projects run in the healthcare sector by other types of aid modalities and thus the amount of non-GBS types of health sector aid should also be controlled for. However, it is likely that adding this control might bias the impact of general budget support for the reason explained above and thus this control is left out and added only to check the robustness of the main results.[10]

[9] Income level, maternal education, and adolescent fertility are all important determinants of child mortality rates. Committing to Child Survival: A Promise Renewed – Progress Report 2014. See more at: http://data.unicef.org/child-mortality/neonatal#sthash.GEfUQnQD.dpuf (accessed 27 October 2010).

[10] Results not reported here but available upon request from the author.

5 Results

Following ADH (2010) and Billmeier and Nannicini (2013) I run the following analysis for each selected GBS-receiving country separately.

(1) First construct a synthetic control using all developing countries[11] available in the control pool (bigger sample size and power of the test).

(2) Then narrow down the group of possible controls using 'common economic zone' to find countries with 'common support' (Billmeier and Nannicini 2013).[12] Thus I am able to choose a synthetic control that follows most closely the pre-treatment outcome trend of the treatment country.

(3) Once the synthetic control is formed, a placebo analysis is run for each country in the pool of controls and the results with our actual case study country.

(4) Compare the root mean squared prediction error pre- and post-treatment year for the actual treatment country and the artificially assigned treatment countries.

(5) Graphically assess the robustness of our results. Calculate the *p*-value to test the randomness of the estimated effect.

5.1 Impact of GBS on health expenditures

Using the synthetic control approach, I construct a convex combination of countries in the control pool that follows as closely as possible the pre-treatment trend of outcome variable for each GBS country.

Table 2 shows the outcome and covariate means as well as the root mean squared prediction error (RMSPE) for steps (1) and (2). The first column shows the actual pre-treatment means for each country in question and columns SC1 and SC2 stand for the synthetic control constructed in steps (1) and (2) correspondingly. SC1 uses all possible control countries and SC2 limits the pool countries from the same region. For each country, the synthetic control with a smaller RMSPE value is chosen for further analysis, which means that SC1 is used for all other countries except for Burkina Faso, Mozambique, and Uganda where SC2 is used.

Figure 8 combines the results for each country. The start of the intervention is marked with a vertical line and it is adjusted for each case study country separately. If a sufficiently accurate pre-period match is found, a constructed synthetic control (dashed line) reveals how the outcome variable would have evolved in the post-treatment period in the absence of treatment.

Countries are divided into three groups according to the results. For countries in Group A, results indicate that health sector expenditures have increased more for the case study country after GBS was launched than for its synthetic control. Out of the top 12 GBS-receiving countries examined, half of them belong (more or less) to Group A, including Tanzania, Burkina Faso, Rwanda, Malawi, and to a certain extent Niger and Burundi. For Group B, the synthetic control approach also more or less worked but the results indicate that there is no traceable

[11] Countries that have not received GBS.

[12] For neonatal mortality rates, low-income countries (LICs) and countries from the same economic zone are included in the more limited pool of 'common support'.

difference between the GBS country and its synthetic control. Group B countries include Mozambique, Ghana and Uganda. Group C contains countries for which the method failed to construct a credible synthetic control for one reason or another. Group C countries are Sierra Leone, Mali, and Madagascar.

Robustness tests are run for each country separately and the results are presented in Appendix Figures A1–A12. In each of these figures, the first panel on the left compares the development in health expenditures in the case study country and its synthetic control. The lower left-hand side panel reports placebo test results where each potential control country is separately assigned to be a 'treatment country', with the rest of the group acting as its potential control countries; the figure shows the outcome difference between each of the treated placebo countries and their synthetic controls. It is also possible to assess the randomness of our results numerically: the figure on the right-hand side shows the relationship between the RMSPE pre- and post-treatment year for the actual treatment country and the artificially assigned treatment countries. The bigger the post-RMSPE/pre-RMSPE for the case study country compared to the placebo effects, the more likely it is that there has been a change in the outcome variable that can be attributed to the treatment. In other words, the effect estimated for the case of interest can be evaluated against the distribution of placebo effects (ADH 2010). If the magnitude of the estimated effect (for the case of interest) falls clearly inside the distribution of placebo effects, it is likely that a large synthetic control estimate for the effect of intervention is just a random coincidence. This probability can be operationalized through the use of *p*-values.

Figures A1–A4 show that, for Tanzania, Burkina Faso, Rwanda and Malawi, receiving budget support has clearly increased government health sector spending per GDP. These countries have, on average, health expenditures per GDP 1.7 percentage points higher after introduction of GBS than they otherwise would have had (*p*-values are listed in Table 3). There are strong indications that the estimated effect of GBS is quite large relative to the distribution of placebo effects for the countries in the control pool. For Niger and Burundi, the results are shown in Figures A5 and A6. They have also increased their health sector spending, and even though the results are not as strong as for the previous four, they nevertheless signal that GBS has had an increasing effect on health sector spending in these countries. Synthetic control country weights for each Group A country are reported in Table 6.

For Mozambique, Ghana, and Uganda (Figures A7–A9) the synthetic control estimate does not reveal any significant increase in health sector spending. The size of their estimated effects falls well inside the distribution of placebo effects thus resulting in p-values greater than 0.2 (Table 3). Country weights for each synthetic control are reported in Table 4.

5.2 Impact of GBS on neonatal mortality rate

Looking only at health sector expenditure does not tell us much about what has been done with the money and whether expanded spending has translated into improved access to or improved quality within the existing healthcare system. While data on access to healthcare services are not widely available, the neonatal mortality rate is used as a proxy. A declining neonatal mortality rate can be seen as an indication of better access to healthcare for mothers and/or of more skilled health personnel (UNICEF 2014; WHO 2010b). Expanded health sector financing, if used appropriately, should have the impact of improving either access or the skill levels of health workers. A situation where government health sector spending increases without any significant impact on the mortality rate should, at the very least, indicate a failure in health policy design.

Table 5 shows the pre-intervention outcome and covariate means for the actual case countries and the synthetic controls 1 and 2 (SC1 and SC2 respectively). SC1 uses all possible control

countries and SC2 limits the control country pool to only low-income countries or countries from the same geographical region. Again choosing the synthetic control with the smallest RMSPE, SC1 is selected for Rwanda, Burundi, Mali and Sierra Leone, and SC2 is used for Tanzania, Malawi, Niger, Burkina Faso, Mozambique, Uganda, and Ghana.

Figure 9 shows the trends in neonatal mortality rate in all the 12 case study countries and the trends of their synthetic controls.[13] I focus the analysis only on countries where there is a (strong) indication that health sector expenditure has increased more than in the synthetic control country after the introduction of GBS. These countries are Malawi, Tanzania, Burkina Faso, Rwanda, Burundi, and Niger. Again, robustness tests for each country are included in the Appendix Figures A13–A24. For Burkina Faso, Malawi, Tanzania, Burundi and, to some extent, Niger, it is clear that something has happened to the neonatal mortality rate parallel with the increase in health sector spending (Appendix Figures A19–A23). In these five countries the decline in the mortality rate has been faster than in their synthetic controls; this is especially the case for Malawi, Burkina Faso, and Tanzania where the neonatal mortality rate is on average 3.8 neonatal deaths lower than it would otherwise have been. In Niger, the effect seems to be a bit smaller and there the synthetic control is less accurate to begin with. Rwanda's mortality rate is strongly influenced by the civil war and massacre of the 1990s and the synthetic control approach fails to match its pattern (Appendix Figure A18). The *p*-values are reported in Table 6 and the country weights for each synthetic control in Table 7.

One GBS-receiving country case where the neonatal mortality has actually declined visibly slower after an introduction of GBS is Ghana (Appendix Figure A20). Despite some increase in health spending (not significant), the Ghanaian government seems not to have been able to translate GBS funding into wider access to or better quality health services. This finding is, in fact, supported by country-specific evaluation reports (IOB 2012; Lawson et al. 2007), which could not report any vast improvement in the scale and quality of healthcare services. By 2010, Ghana still did not have a national healthcare system in place, with the poorest people having to pay for their personal healthcare, thus limiting access even further.

6 Conclusion

The pressure to prove what has been achieved through development aid grows stronger as donor country economies face challenges of their own. In addition, corruption scandals and electoral frauds have led donors to suspend general budget support on several occasions in recent years. Despite demands for greater transparency and accountability in donor countries, the non-earmarked nature of GBS funds makes rigorous measuring of outcomes and impacts fairly difficult. While country-specific evaluations and cross-country estimations have both been conducted in order to track down the impact of GBS, they typically fail to reveal counterfactuals for individual countries.

The synthetic control method has been described as 'a bridge between a case study approach and a cross-country econometric response' (Billmeier and Nannicini 2013) since it allows us to look

[13] One of the limitations of the synthetic control method is demonstrated in the case of Mozambique (Appendix Figure A18) where the method fails to create a matching synthetic control for the pre-treatment period. The neonatal mortality rate in Mozambique is much higher at the start of the 1990s than in any of the potential control countries. Since the synthetic control method operates with positive weights to avoid extrapolation, it is not possible to construct a matching synthetic control beyond the support of comparison units.

at each country individually and to create a counterfactual to reveal if the observed changes in outcome variables would also have happened in the absence of the instrument.

The results indicate that, at least in Tanzania, Burkina Faso, Rwanda and Malawi—and to some extent in Niger and Burundi—receiving GBS has had a positive effect on government health sector spending. Further, parallel to the increased health sector expenditures, neonatal mortality rates declined, especially in Malawi, Tanzania and Burkina Faso, relatively more than in their synthetic control countries. The case of Burkina Faso is interesting, since while it remains one of least developed countries in the world, it has managed to increase its health sector budget significantly (more than doubling per capita values since the 1990s). The government of Burkina Faso has been, for example, providing an 80 per cent subsidy towards the cost of assisted delivery and emergency obstetric care that has enabled the rate of skilled birth attendance or supervised delivery to rise steadily (De Alegri et al. 2012; UNICEF 2010). This, in turn, has most likely helped to reduce neonatal and maternal mortality rates (EC 2013; WHO 2010c).

In Mozambique however—a major recipient of budget support—the results indicate that the money has not been channelled into the health sector to any significant degree. A likely reason for that might be that the Mozambique health sector receives substantial off-budget funding (e.g. from the PROSAUDE-programme). Therefore, GBS funding was not allocated to the health sector in substantial quantities, but was targeted instead to the education sector (Lawson et al. 2014). This could also explain the faster-than-average decline in the neonatal mortality rate; this had already started in the mid-1990s and cannot be attributed to receiving general budget support.

One caveat of using the synthetic control method in the development context is that it cannot isolate the impact of a particular reform if the country is reforming in a large number of areas at once (McKenzie 2013). However, a larger reform is, in effect, inbuilt in receiving general budget support: political dialogue around general budget support is facilitated by Poverty Reduction Budget Papers that describe recipient governments' wider approach to reducing poverty in different sectors. Thus general budget support funds can be seen as additional financing, enabling a number of reforms at the same time.

One major disadvantage of the setting used in this paper is that data for government health sector spending only extend back to 1995, leaving only a few years for pre-treatment period matching. It might make sense to look at other possible outcome variables (in the health sector or in the education sector) with longer datasets and examine whether results indicate any significant impact of GBS when the time period is expanded.

References

Abadie, A., A. Diamond, and J. Hainmueller (2010). 'Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. Applications and Case Studies'. *Journal of the American Statistical Association*, 105(490): 493–505.

Abadie, A., A. Diamond, and J. Hainmueller (2015). 'Comparative Politics and the Synthetic Control Method'. *American Journal of Political Science*, 59(2): 495–515.

Abadie, A., and J. Gardeazabal (2003). 'The Economic Costs of Conflict: A Case Study of the Basque Country'. *American Economic Review*, 93(1): 113–32.

Alavuotunki, K. (2015). 'General Budget Support – Impact on Government Expenditure Allocation'. Mimeo.

Alavuotunki, K., and S. Sandström (2015). 'General Budget Support as an Aid Instrument – Allocation and Impact'. Mimeo.

Antunes, A.F., K. Xu, C.D. James, P. Saksena, N. van de Maele, G. Carrin, and D.B. Evans (2013). 'General Budget Support. Has it Benefited the Health Sector?'. *Health Economics*, 22: 1440–51

Beynon, J., and A. Dusu (2010). 'Budget Support and MDG Performance'. Development Paper 2010/01. Brussels: European Commission.

Bhutta, Z.A., M. Chopra, H. Axelson, P. Berman, T. Boerma, J. Bryce, F. Bustreo, E. Cavagnero, G. Cometto, B. Daelmans, A. de Francisco, H. Fogstad, H. Gupta, L. Laski, J. Lawn, B. Maliqi, E. Mason, C. Pitt, J. Requejo, A. Starrs, C.G. Victora, and T. Wardlaw (2010). 'Countdown to 2015 Decade Report (2000–10): Taking Stock of Maternal, Newborn, and Child Survival'. *Lancet*, 375: 2032–44.

Bigsten, A., J. Platteau, and S. Tengstam (2011). *The Aid Effectiveness Agenda: The Benefits of Going Ahead*. Gothenburg: SOGES, European Commission.

Billmeier, A., and T. Nannicini (2013). 'Assessing Economic Liberalization Episodes. A Synthetic Control Approach'. *Review of Economics and Statistics*, 95(3): 983–1001.

Caputo, E., A. de Kemp, and A. Lawson (2011). 'Assessing the Impacts of Budget Support – Case Studies in Mali, Tunisia and Zambia'. *Evaluation Insights*, 2. Paris: European Commission.

Caputo, E., C. Valmarana, S. Smail, D. Chiwandamira, M. Smith, and J. Vanheukelom (2013). 'Evaluation of Budget Support in South Africa. Final Report Volume 1'. Independent Evaluation Jointly Managed by the European Commission and the Government of South Africa. Freiburg,: Particip.

Collier, P., P. Guillaumont, S. Guillaumont, and J.W. Gunning (1997). 'Redesigning Conditionality'. *World Development*, 25(9): 1399–407.

De Allegri, M., V. Ridde, V.R. Louis, M. Sarker, J. Tendrebogo, M.Y.O. Mueller, and A. Jahn (2012). 'The Impact of Targeted Subsidies for Facility 36-based Delivery on Access to Care and Equity – Evidence from a Population-based Study in Rural Burkina Faso'. *Journal of Public Health Policy*, 33: 439–53.

de Kemp, A., J. Faust, and S. Leiderer (2011). 'Between High Expectations and Reality: An Evaluation of Budget Support in Zambia'. IOB Evaluation 356. Bonn/The Hague/ Stockholm: BMZ/IOB/Sida.

Dijkstra, G. (2013). 'Governance or Poverty Reduction? Assessing Budget Support in Nicaragua'. *Journal of Development Studies*, 49(1): 110–24.

Dijkstra, G., and A. de Kemp (2015). 'Challenges in Evaluating Budget Support and How to Solve Them'. *Evaluation*, 21(1): 83–98.

Dollar, D., and J. Svensson (2000). 'What Explains the Success or Failure of Structural Adjustment Programs?'. *Economic Journal*, 110: 894–917.

EC (2013). 'European Commission Press Release 2013'. Available at: http://europa.eu/rapid/press-release_IP-13-1031_en.htm (accessed 25 October 2015).

Elbers, C., J.W. Gunning, and K. de Hoop (2009). 'Assessing Sector-wide Programs with Statistical Impact Evaluation: A Methodological Proposal'. *World Development*, 37: 513–20.

Fosu, A.K. (2010). 'The External Debt-servicing Constraint and Public Expenditure Composition in Sub-Saharan Africa'. *African Development Review*, 22(3): 378–93.

Fosu, A.K., and M. Quattri (2012). 'On the Impact of External Debt and Aid on Public Expenditure Allocation in Sub-Saharan Africa after the Launch of the HIPC Initiative'. Working Paper 2012/42. Helsinki: UNU-WIDER.

Gathani, S., M. Santini, and D. Stoelinga (2013). 'Innovative Techniques to Evaluate the Impact of Private Sector Development Reforms: An Application to Rwanda and 11 Other Countries'. Mimeo.

IDD and Associates (2006). 'Joint Evaluation of General Budget Support: Synthesis Report'. Study undertaken by IDD and Associates for the OECD. Birmingham: University of Birmingham.

IOB (2012). 'Budget Support: Conditional Results: Review of an Instrument, 2000–2011'. IOB Evaluation 369. The Hague: Ministry of Foreign Affairs, Policy and Operations Evaluation Department.

Kirkpatrick, A.J., and L.S. Bennear (2014). 'Promoting Clean Energy Investment: An Empirical Analysis of Property-assessed Clean Energy'. *Journal of Environmental Economics and Management*, 68: 357–75.

Lawn, J.E., D. Osri, A. Adler, and S. Cousens (2008). 'Four Million Neonatal Deaths: Counting and Attribution of Cause of Death'. *Paediatric and Perinatal Epidemiology*, 22: 410–16.

Lawson, A., M. Bartholomew, A.M. Bibi, et al. (2014). 'Independent Evaluation of Budget Support in Mozambique'. Hove: Itad. Available at: www.oecd.org/derec/ec/Mz-BS-Eval-Final-Report-Vol-1.pdf (accessed 25 October 2015).

Lawson, A., G. Boadi, A. Ghartey, et al. (2007). 'Joint Evaluation of Multi-donor Budget Support to Ghana: Evaluation of Outputs, Outcomes and Impacts and Recommendations on Future Design and Management of Ghana MDBS'. London: Overseas Development Institute. Available at: www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/247.pdf (accessed 25 October 2015).

Lawson, A., J. Habas, M. Keita, E. Paul, B. Versailles, and A. Murray-Zmijewski (2011). 'Joint Evaluation of Budget Support Operations in Mali, 2003–2009 (Executive Summary)'. Brussels: Joint Evaluation Unit.

Lawson, A., J. Kipokola, T. Ruffer, P. Tidemand, E. Paul, and A. Bartholomew (2013). 'Joint Evaluation of Budget Support to Tanzania: Lessons Learned and Recommendations for the Future'. Independent evaluation jointly managed by the European Commission, the Ministry of Foreign Affairs of Denmark, Irish Aid, the Ministry of Foreign Affairs of the Netherlands and the Ministry of Finance of Tanzania.

Lépine, A., M. Lagarde, and A. Le Nestour (2015). 'Free Primary Care in Zambia: An Impact Evaluation Using a Pooled Synthetic Control Method'. Mimeo.

Łukasz, M. (2014). 'The Causal Links between Aid and Government Expenditures'. Tinbergen Institute Discussion Paper TI 2014-012/V. Amsterdam: Tinbergen Institute.

McKenzie, D. (2013). 'Evaluating Regulatory Reforms using the Synthetic Control Method. Development Impact'. World Bank online blog. Available at: http://blogs.worldbank.org/impactevaluations/evaluating-regulatory-reforms-using-the-synthetic-control-method (accessed 25 October 2015).

Mideksa, T.K. (2013). 'The Economic Impact of Natural Resources'. *Journal of Environmental Economics and Management*, 65: 277–89.

Molenaers, N., L. Cepinskas, and B. Jacobs (2010). 'Budget Support and Policy/Political Dialogue: Donors' Practices in Handling (Political) Crises'. Discussion Paper 2010–06. Antwerp: University of Antwerp/IOB.

Singhal, S., and R. Nilakantan (2012). 'Naxalite Insurgency and the Economic Benefits of a Unique Robust Security Response'. Mimeo.Rajaratnam, J.K., J.R. Marcus, A.D. Flaxman, H. Wang, A. Levin-Rector, L. Dwyer, M. Costa, A.D. Lopez, and C.J.L. Murray (2010). 'Neonatal, Postneonatal, Childhood, and Under-5 Mortality for 187 Countries, 1970–2010: A Systematic Analysis of Progress towards Millennium Development Goal 4'. *Lancet*, 375(9730): 1988–2008.

UNICEF (2009). 'State of the World for Children 2009. Maternal and Newborn Health'. Available at: www.unicef.org/sowc09/report/report.php (accessed 25 October 2015).

UNICEF (2010). 'Annual Report for Burkina Faso'. Available at: www.unicef.org/about/annualreport/files/Burkina_Faso_COAR_2010.pdf (accessed 25 October 2015).

UNICEF (2011). 'Levels and Trends in Child Mortality. Report 2011'. Available at: www.unicef.org/media/files/Child_Mortality_Report_2011_Final.pdf (accessed 25 October 2015).

UNICEF (2012). *Issue Briefs: Maternal and Child Health*. Jakarta: UNICEF.

UNICEF (2013). 'Committing to Childs' Survival: A Promise Renewed – Progress Report 2013'. Available at: www.unicef.org/publications/index_70354.html (accessed 3 November 2015).

UNICEF (2014). 'Committing to Childs' Survival: A Promise Renewed – Progress Report 2014'. Available at: http://data.unicef.org/child-mortality/neonatal#sthash.GEfUQnQD.dpuf (accessed 25 October 2015).

WHO (2010a). 'General Budget Support – Has it Benefited the Health Sector?' World Health Report. Background Paper 14. Available at: www.who.int/healthsystems/topics/financing/healthreport/whr_background/en (accessed 3 November 2015).

WHO (2010b). 'Countdown to 2015. Decade Report (2000–2010) with Country Profiles. Taking Stock of Maternal, Newborn and Child Survival'. Available at: www.who.int/maternal_child_adolescent/documents/9789241599573/en/ (accessed 3 November 2015).

WHO (2010c). 'Maternal and Perinatal Health Profile – Burkina Faso'. Available at: www.who.int/maternal_child_adolescent/.../profiles/maternal/bfa.pdf (accessed 3 November 2015).

WDI (2014). *World Bank Development Indicators*. Washington, DC: World Bank.

Tables and figures

Table 1: Top GBS-receiving countries in 2005

Country	GBS of gov't exp (%)	GBS of total ODA (%)	GBS per GDP (%)	GBS in USD (million)
Rwanda	38	52	7	188
Sierra Leone	32	37	5	78
Mozambique	26	35	5	321
Tanzania	17	55	4	552
Malawi	17	25	4	98
Burundi	15	18	4	46
Uganda	13	30	3	229
Madagascar	12	32	2	105
Niger	11	28	3	86
Burkina Faso	10	37	3	162
Ghana	10	46	3	285
Mali	9	17	2	83

Source: OECD-DAC CRS database and author's own calculations for general budget support; Net Aid Transfers database by Roodman for total Overseas Development Assistance; World Bank, World Development Indicators for government expenditures and GDP.

Table 2: Covariates and outcome means[14]

	TAZ	SC1	SC2		MOZ	SC1	SC2
Health exp. of GDP	1.36	1.35	1.36	Health exp. of GDP	3.10	3.10	3.08
Log GDP per capita	5.64	5.77	5.96	Log GDP per capita	5.35	7.70	7.32
Log depend. rate	4.52	4.51	4.52	Log depend. rate	4.48	4.38	4.48
Military exp. of GDP	1.47	1.53	1.47	Military exp. of GDP	1.21	1.22	3.28
Log population	17.28	15.88	15.60	Log population	16.64	16.64	16.05
RMSPE		0.12	0.12	RMSPE		0.26	0.19
	RWA	SC1	SC2		GHA	SC1	SC2
Health exp. of GDP	1.99	1.97	2.00	Health exp. of GDP	2.82	2.82	2.77
Log GDP per capita	5.33	5.67	6.20	Log GDP per capita	6.06	6.79	6.95
Log depend. rate	4.49	4.41	4.50	Log depend. rate	4.40	4.40	4.40
Military exp. of GDP	4.54	4.53	4.54	Military exp. of GDP	0.75	1.17	1.98
Log population	15.70	16.63	15.70	Log population	16.70	16.15	15.53
RMSPE		0.13	0.31	RMSPE		0.27	0.39
	MWI	SC1	SC2		MDG	SC1	SC2
Health exp. of GDP	1.78	1.77	1.78	Health exp. of GDP	2.06	2.05	2.06
Log GDP per capita	5.39	7.13	5.98	Log GDP per capita	5.64	7.10	7.04
Log depend. rate	4.53	4.52	4.57	Log depend. rate	4.53	4.53	4.51
Military exp. of GDP	0.86	0.91	1.95	Military exp. of GDP	1.26	1.61	1.27
Log population	16.16	16.16	16.16	Log population	16.51	16.35	16.41
RMSPE		0.18	0.14	RMSPE		0.26	0.29
	NER	SC1	SC2		MLI	SC1	SC2
Health exp. of GDP	1.53	1.53	1.47	Health exp. of GDP	2.47	2.45	2.46
Log GDP per capita	5.58	6.20	5.88	Log GDP per capita	5.90	5.99	7.22
Log depend. rate	4.62	4.55	4.58	Log depend. rate	4.59	4.40	4.56
Military exp. of GDP	1.06	1.88	11.92	Military exp. of GDP	1.65	1.77	2.92
Log population	16.14	16.16	14.57	Log population	16.09	16.68	14.85
RMSPE		0.10	0.17	RMSPE		0.34	0.39
	BFA	SC1	SC2		SLE	SC1	SC2
Health exp.of GDP	1.88	1.88	1.88	Health exp.of GDP	2.63	2.63	2.60
Log GDP per capita	5.79	6.01	6.53	Log GDP per capita	5.63	6.31	6.39
Log Depend. rate	4.60	4.60	4.48	Log Depend. rate	4.44	4.44	4.51
Military exp. of GDP	1.16	11.15	1.28	Military exp. of GDP	2.41	2.61	19.06
Log population	16.21	16.01	16.22	Log population	15.20	15.21	15.18
RMSPE		0.11	0.09	RMSPE		0.64	0.70
	BDI	SC1	SC2		UGA	SC1	SC2
Health exp. of GDP	1.85	1.85	1.82	Health exp. of GDP	1.77	1.76	1.77
Log GDP per capita	5.05	6.07	6.34	Log GDP per capita	5.53	6.16	6.21
Log depend. rate	4.68	4.54	4.55	Log depend. rate	4.67	4.53	4.49
Military exp. of GDP	5.70	5.71	5.71	Military exp. of GDP	2.45	2.46	2.46
Log population	15.69	15.69	14.72	Log population	16.91	16.89	16.17
RMSPE		0.24	0.28	RMSPE		0.16	0.11

Notes: BDI = Burundi; BFA = Burkina Faso; GHA = Ghana; Health exp. of GDP = Health expenditure as a percentage of GDP; MDG = Madagascar; MLI = Mali; MOZ = Mozambique; MWI = Malawi; NER = Niger; RWA = Rwanda; SLE = Sierra Leone; TAZ = Tanzania; UGA = Uganda.

Source: Author's calculations.

[14] Synthetic Control 1 (SC1) uses data from all the potential control countries while synthetic control 2 (SC2) limits the selection to countries from the same economic region (SSA).

Table 3: Probability of obtaining an estimate at least as large as the one obtained for the unit representing the case of interest when the intervention is reassigned at random in the dataset

Country	*p*-value
Tanzania	1/54 ≈ 0.019
Rwanda	1/50 = 0.020
Burkina Faso	1/15 ≈ 0.067
Malawi	1/50 ≈ 0.019
Niger	4/54 ≈ 0.072
Burundi	4/54 ≈ 0.072
Ghana	14/54 ≈ 0.255
Mozambique	6/13 ≈ 0.461
Uganda	9/13 ≈ 0.69

Source: Author's calculations.

Table 4: Country weights; health expenditures

Tanzania (SC1)	Guinea 0.67 Gambia 0.16 Uzbekistan 0.11 Yemen 0.06	Niger (SC1)	Gambia 0.31 Guinea 0.19 Nigeria 0.23 Uzbekistan 0.07 Yemen 0.21
Rwanda (SC1)	Eritrea 0.13 Nepal 0.73 Kyrgyzstan 0.14	Burkina Faso (SC2)	Gambia 0.21 Guinea 0.22 Namibia 0.20 Nigeria 0.38 Swaziland 0.001
Malawi (SC1)	Guatemala 0.78 Gambia 0.10 Swaziland 0.07 Panama 0.05	Burundi (SC1)	Eritrea 0.13 Gambia 0.18 Guinea 0.24 Uzbekistan 0.16 Yemen 0.29
Mozambique (SC2)	Angola 0.32 Namibia 0.43 Nigeria 0.14	Ghana (SC1)	Gambia 0.36 Panama 0.22 Uzbekistan 0.42
Uganda (SC2)	Angola 0.02 Eritrea 0.04 Guinea 0.57 Namibia 0.16 Nigeria 0.21		

Source: Author's calculations.

Table 5: Covariates and outcome means (neonatal mortality rate)[15]

	TAZ	SC1	SC2		MOZ	SC1	SC2
Neonatal mortality rate	40.72	40.67	40.70	Neonatal mortality rate	52.75	52.64	52.51
Adolescent fertility	136.35	138.85	122.99	Adolescent fertility	119.52	194.76	206.11
Log GDP per capita	5.64	6.63	6.03	Log GDP per capita	5.29	6.76	6.63
Population under 15	3.81	3.73	3.82	Population under 15	3.80	3.83	3.84
Female:male in prim. edn.	98.46	92.04	84.43	Female:male in prim. edn.	73.82	81.38	76.16
RMSPE		1.07	0.34	RMSPE		2.31	2.09
	RWA	SC1	SC2		GHA	SC1	SC2
Neonatal mortality rate	42.76	42.70	42.73	Neonatal mortality rate	36.64	36.60	36.66
Adolescent fertility	57.11	119.83	113.44	Adolescent fertility	93.32	107.06	93.91
Log GDP per capita	5.37	6.87	6.36	Log GDP per capita	6.02	6.21	6.69
Population under 15	3.83	3.68	3.78	Population under 15	3.75	3.72	3.75
Female:male in prim. edn.	97.56	93.49	86.76	Female:male in prim. edn.	89.31	89.22	89.25
RMSPE		3.01	3.76	RMSPE		0.46	0.11
	MWI	SC1	SC2		MDG	SC1	SC2
Neonatal mortality rate	45.58	45.54	45.57	Neonatal mortality rate	36.29	36.29	36.24
Adolescent fertility	162.79	158.16	161.35	Adolescent fertility	151.54	128.65	121.24
Log GDP per capita	5.33	6.36	5.59	Log GDP per capita	5.65	6.19	5.85
Population under 15	3.81	3.77	3.79	Population under 15	3.81	3.79	3.81
Female:male in prim. edn.	92.04	82.67	59.33	Female:male in prim. edn.	96.47	89.63	83.24
RMSPE		0.95	0.26	RMSPE		1.15	0.51
	NER	SC1	SC2		MLI	SC1	SC2
Neonatal mortality rate	45.76	45.78	45.73	Neonatal mortality rate	56.89	53.91	53.91
Adolescent fertility	222.00	172.58	161.50	Adolescent fertility	189.88	217.98	217.98
Log GDP per capita	5.61	5.99	5.84	Log GDP per capita	5.86	7.12	7.12
Population under 15	3.87	3.80	3.81	Population under 15	3.83	3.86	3.86
Female:male in prim. edn.	63.21	62.97	61.91	Female:male in prim. edn.	67.10	87.84	87.84
RMSPE		0.29	0.17	RMSPE		3.13	3.13
	BFA	SC1	SC2		SLE	SC1	SC2
Neonatal mortality rate	39.74	39.74	39.75	Neonatal mortality rate	55.79	54.00	54.00
Adolescent fertility	142.88	141.16	137.58	Adolescent fertility	130.82	219.00	219.00
Log GDP per capita	5.72	5.76	6.59	Log GDP per capita	5.71	7.11	7.11
Population under 15	3.85	3.84	3.77	Population under 15	3.78	3.86	3.86
Female:male in prim. edn.	67.15	67.33	90.89	Female:male in prim. edn.	0.73	0.82	0.82
RMSPE		2.27	0.08	RMSPE		1.99	1.99
	BDI	SC1	SC2		UGA	SC1	SC2
Neonatal mortality rate	44.17	44.15	44.15	Neonatal mortality rate	38.04	38.03	38.09
Adolescence fertility	45.60	93.84	159.67	Adolescence fertility	194.10	147.82	154.49
Log GDP per capita	5.17	6.67	6.01	Log GDP per capita	5.43	7.24	7.80
Population under 15	3.89	3.75	3.83	Population under 15	3.88	3.75	3.79
Female:male in prim. edn.	80.83	80.24	74.10	Female:male in prim. edn.	85.42	88.81	85.78
RMSPE		0.38	1.06	RMSPE		0.13	0.11

Source: Author's calculations.

[15] Synthetic Control 1 (SC1) uses from all the potential control countries while Synthetic Control 2 (SC2) limits the selection to countries from the same economic region (SSA).

Table 6: Probability of obtaining an estimate at least as large as the one obtained for the unit representing the case of interest when the intervention is reassigned at random in the dataset

Country	*p*-value
Burkina Faso	1/17 ≈ 0.06
Malawi	1/17≈ 0.06
Burundi	2/48 ≈ 0.04
Tanzania	1/17 ≈ 0.06
Niger	2/17 ≈ 0.12
Rwanda	26/48 ≈ 0.54

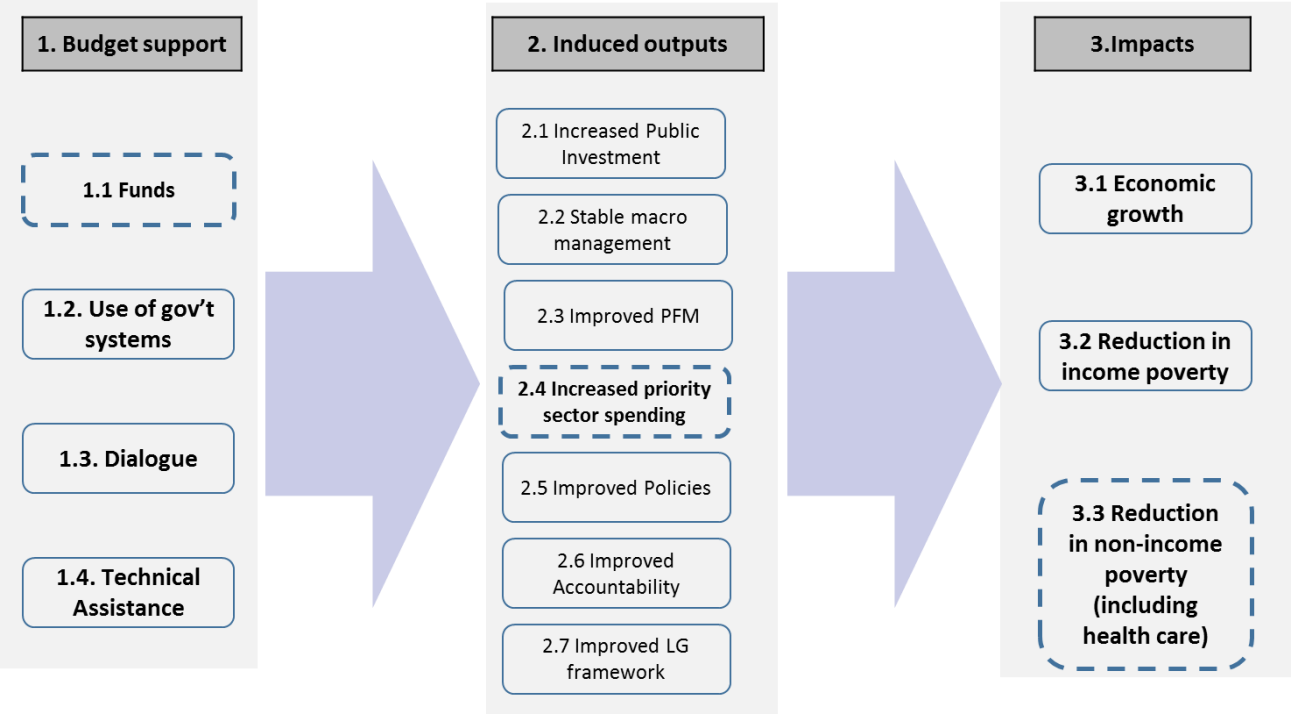
Source: Author's calculations.

Table 7: Country weights; neonatal mortality

Tanzania	Angola 0.487	Niger	Angola 0.01
	Azerbaijan 0.06		Gambia 0.28
	Kazakhstan 0.43		Guinea 0.63
	Nigeria 0.03		Yemen 0.09
Rwanda		Burkina Faso	Angola 0.40
	Not relevant		Nigeria 0.21
			Uzbekistan 0.39
Malawi	Eritrea 0.14	Burundi	Djibouti 0.40
	Nigeria 0.41		Nigeria 0.54
	Nepal 0.45		Sri Lanka 0.06

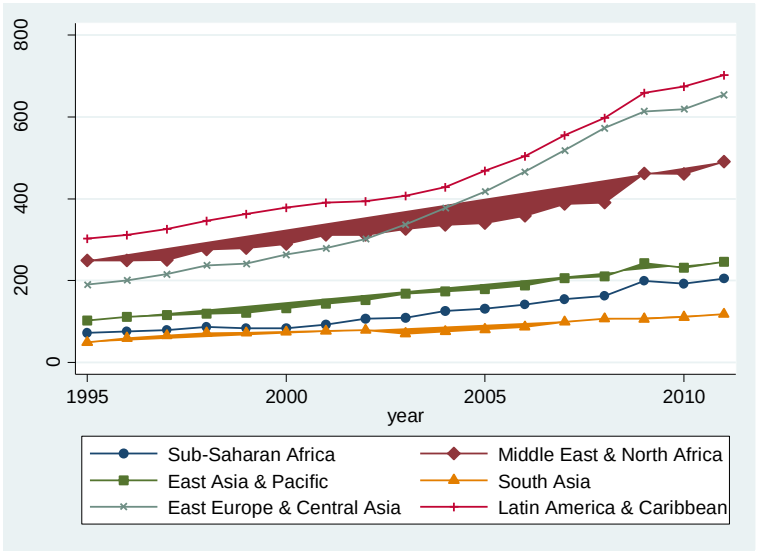
Source: Author's calculations.

Figure 1: Adapted logical framework from joint budget support evaluations



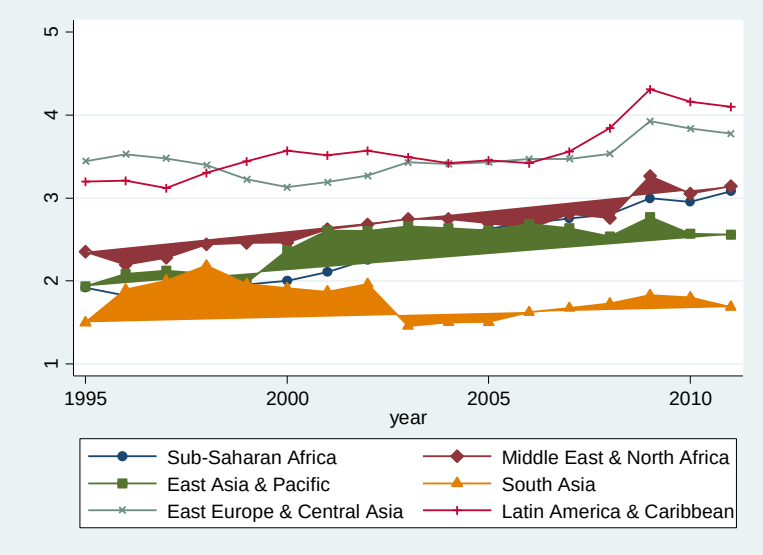
Source: EC 2013: ix (Tanzania evaluation).

Figure 2: Development of government health expenditure per capita by regions



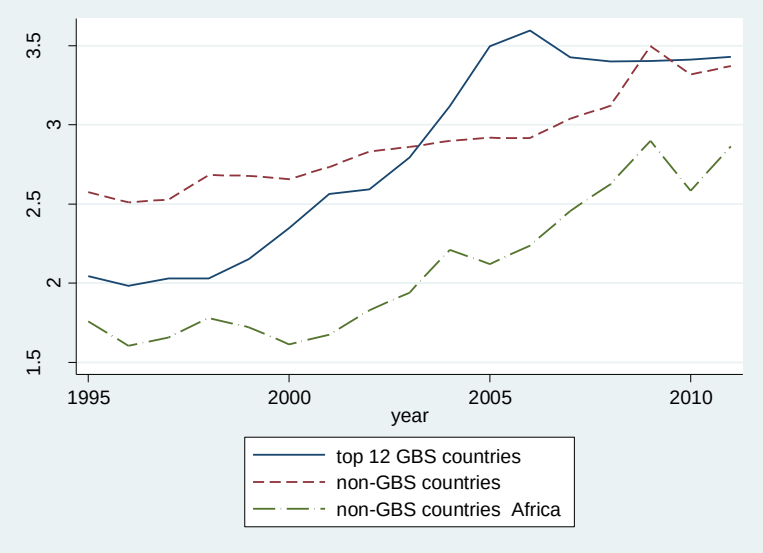
Source: World Development Indicators, the World Bank. Provided by World Health Organization National Health Account database.

Figure 3: Development of government health expenditure per GDP by regions



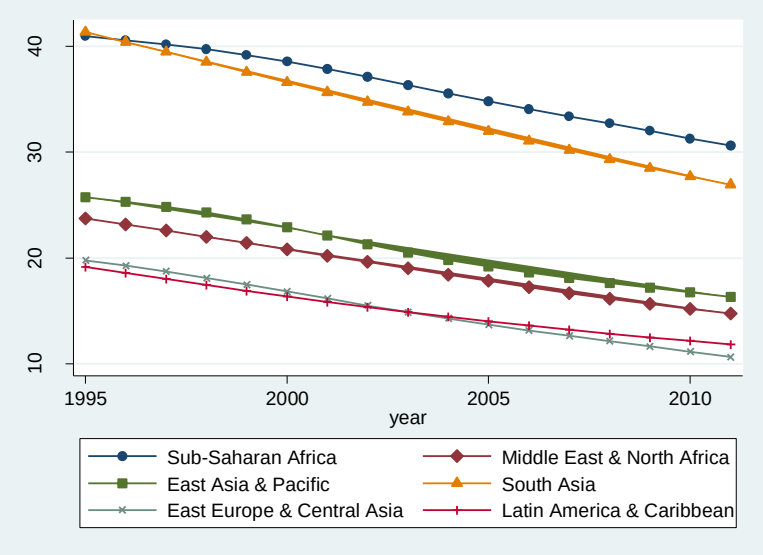
Source: World Development Indicators, the World Bank.

Figure 4: Health sector expenditure per GDP in GBS-receiving and non-GBS-receiving countries



Source: World Development Indicators, the World Bank.

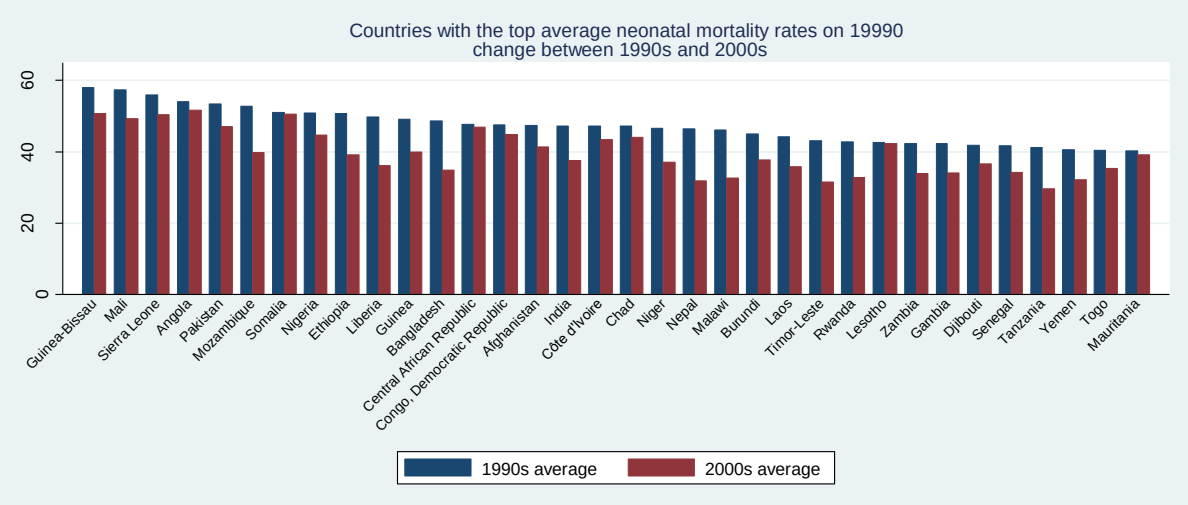
Figure 5: Neonatal mortality rates by region.



Notes: The infant mortality rate is the number of infants dying before reaching one year of age, per 1,000 live births in a given year.

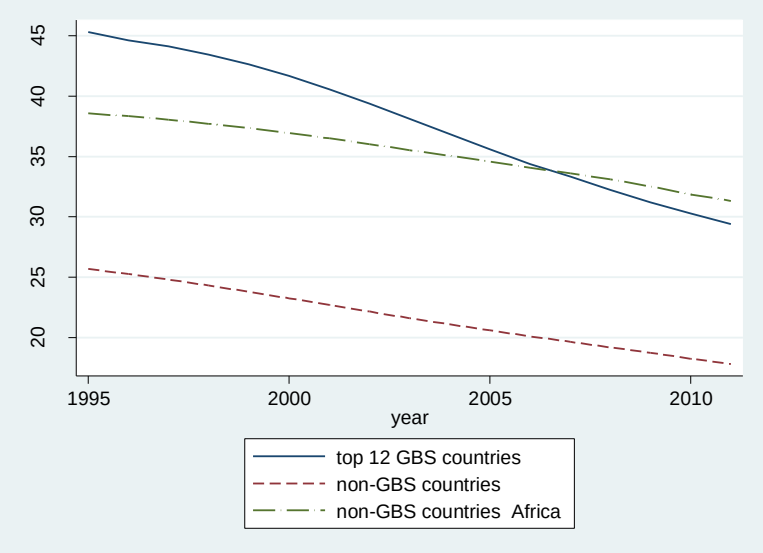
Source: World Development Indicators, the World Bank.

Figure 6: Change in neonatal mortality rate between 1990s and 2000s



Source: World Development Indicators, the World Bank.

Figure 7: Trends in neonatal mortality in GBS-receiving and non-GBS-receiving countries

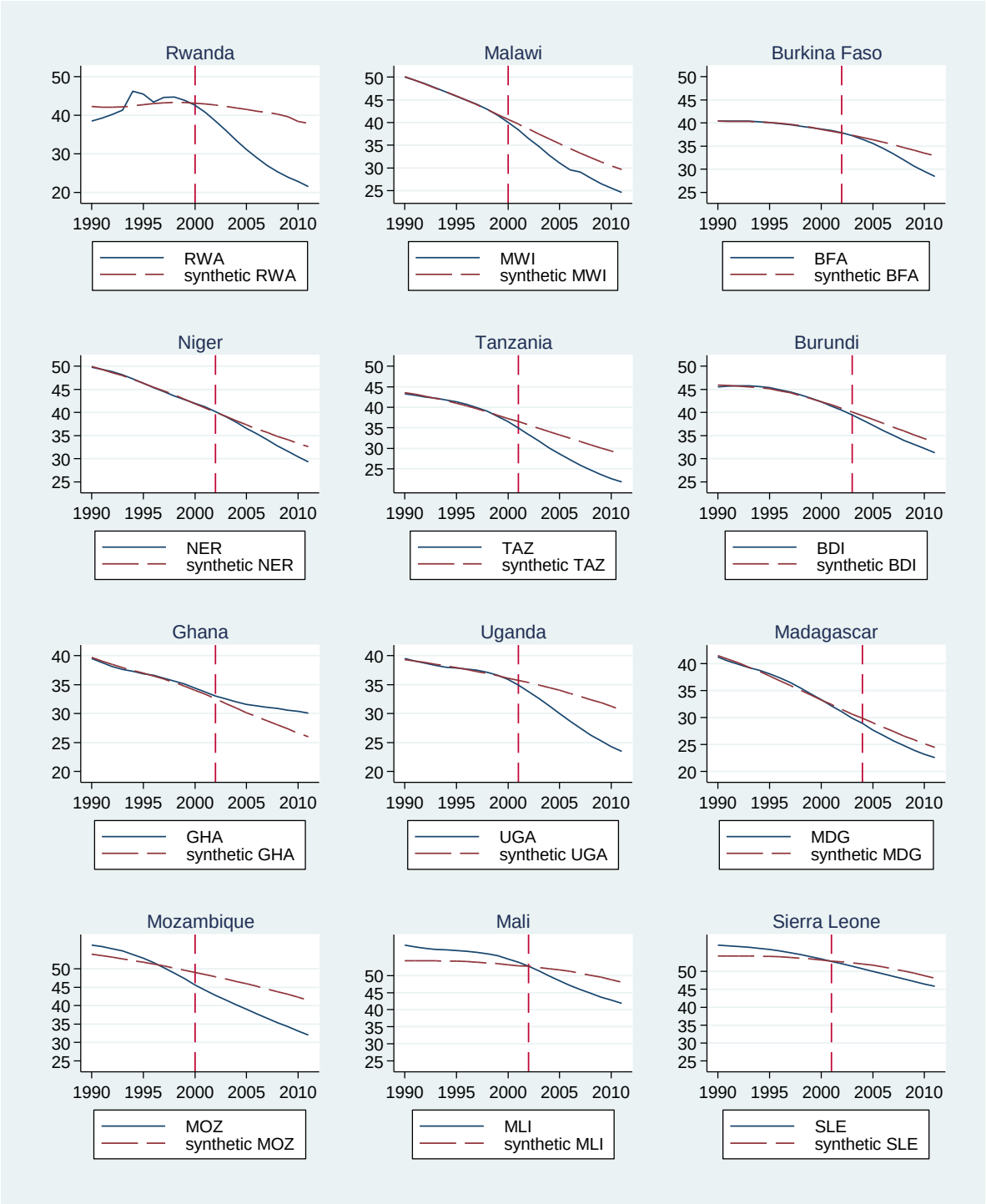


Source: World Development Indicators, the World Bank.

Figure 8: Health expenditure per GDP (%) trends: treated countries vs synthetic controls

Source: Author's calculations.

Figure 9: Trends in neonatal mortality rate: treated countries vs synthetic controls



Source: Author's calculations.

APPENDIX

Table A1: The start of GBS as an aid modality

SSA		**South and East Asia and Pacific**		**Latin America and Caribbean**	
Benin (Dahomey)	Early 2000s	Bangladesh	Early 2000s	Bolivia	Early 2000s
Burkina Faso	Early 2000s	Indonesia	Early 2000s	Haiti	Early 2000s
Burundi	Early 2000s	Laos	Early 2000s	Nicaragua	Early 2000s
Cameroon	Early 2000s	Pakistan	Early 2000s	Guyana	Mid-2000s
Cape Verde	Early 2000s	Timor Leste	Early 2000s	Dominican Republic	Late 2000s
Central African Rep.	Early 2000s	Viet Nam	Early 2000s	Honduras	Late 2000s
Ethiopia	Early 2000s	Bhutan	Late 2000s	Paraguay	Late 2000s
Ghana	Early 2000s	Cambodia	Late 2000s	Argentina	no GBS
Lesotho	Early 2000s	India	Late 2000s	Brazil	no GBS
Malawi	Early 2000s	Solomon Islands	Late 2000s	Chile	no GBS
Mali	Early 2000s	China	no GBS	Colombia	no GBS
Mauritania	Early 2000s	Fiji	no GBS	Costa Rica	no GBS
Mozambique	Early 2000s	Korea, North	no GBS	Cuba	no GBS
Niger	Early 2000s	Malaysia	no GBS	Ecuador	no GBS
Rwanda	Early 2000s	Mongolia	no GBS	El Salvador	no GBS
Senegal	Early 2000s	Nepal	no GBS	Guatemala	no GBS
Sierra Leone	Early 2000s	Oman	no GBS	Mexico	no GBS
Tanzania	Early 2000s	Papua New Guinea	no GBS	Panama	no GBS
Uganda	Early 2000s	Philippines	no GBS	Peru	no GBS
Zambia	Early 2000s	Sri Lanka	no GBS	Suriname	no GBS
Kenya	Mid-2000s	Thailand	no GBS	Trinidad and Tobago	no GBS
Madagascar	Mid-2000s			Uruguay	no GBS
Togo	Mid-2000s			Venezuela	no GBS
Chad	Late 2000s	**Central Asia and Europe**			
Comoros	Late 2000s	Armenia	Early 2000s	**North Africa, Middle East**	
Congo, Dem. Rep	Late 2000s	Georgia	Early 2000s	Afghanistan	Early 2000s
Guinea-Bissau	Late 2000s	Tajikistan	Early 2000s	Palestinian Territory	Early 2000s
Mauritius	Late 2000s	Albania	Mid-2000s	Tunisia	Early 2000s
Angola	no GBS	Bosnia-Herzegovina	Mid-2000s	Jordan	Mid-2000s
Botswana	no GBS	Moldova	Late 2000s	Morocco	Late 2000s
Congo, Republic	no GBS	Kosovo	Late 2000s	Algeria	no GBS
Côte d'Ivoire	no GBS	Serbia & Montenegro	Late 2000s	Djibouti	no GBS
Equatorial Guinea	no GBS	Azerbaijan	no GBS	Egypt	no GBS
Eritrea	no GBS	Belarus	no GBS	Iraq	no GBS
Gabon	no GBS	Croatia	no GBS	Iran	no GBS
Gambia	no GBS	Kazakhstan	no GBS	Lebanon	no GBS
Guinea	no GBS	Kyrgyzstan	no GBS	Libya	no GBS
Namibia	no GBS	Macedonia	no GBS	Saudi Arabia	no GBS
Nigeria	no GBS	Montenegro	no GBS	Yemen	no GBS
Somalia	no GBS	Myanmar	no GBS		
South Africa	no GBS	Turkey	no GBS		
Sudan	no GBS	Turkmenistan	no GBS		
Swaziland	no GBS	Ukraine	no GBS		
Zimbabwe	no GBS	Uzbekistan	no GBS		

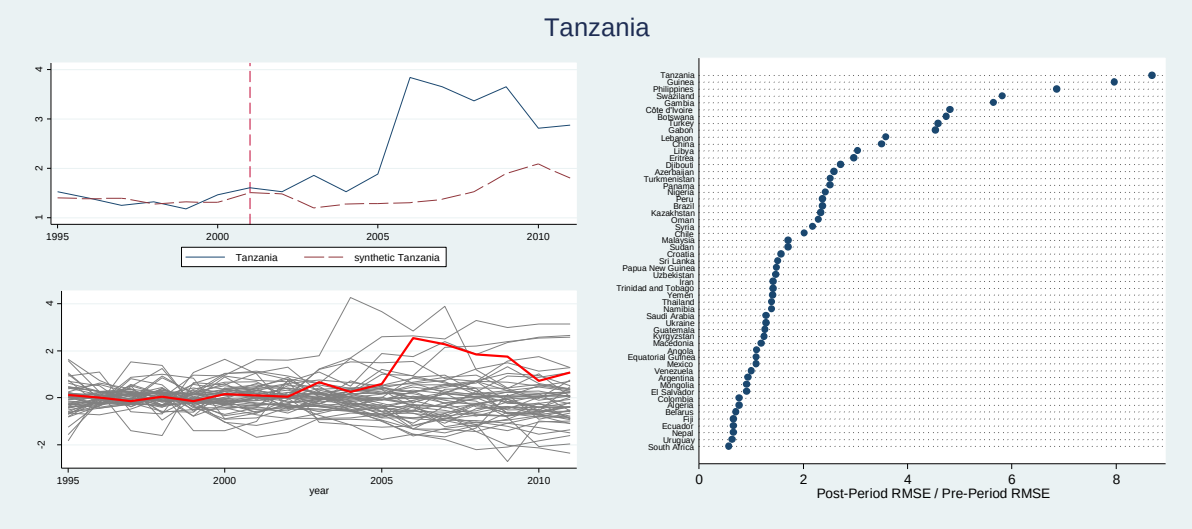
Source: Several sources containing information on the timing of introduction of GBS, including country evaluation reports and donor reports (combined by the author).

Table A2: Data description and sources

Variable	Description	Source
General budget support	The definition of GBS used follows the standard definition used by OECD-DAC: (1) non-earmarked financial support directly to the recipient country's budget and (2) should be part of a co-ordinated and harmonized donor agenda. Any actions related to debt, sector budget support or ad hoc support disbursed under the heading of GBS but given by only one single bilateral donor are excluded.	OECD-DAC CRS database and author's own edits for general budget support based on different evaluation and country reports.
Health expenditure per GDP	Total health expenditure is the sum of public and private health expenditure. It covers the provision of health services (preventive and curative), family planning activities, nutrition activities, and emergency aid designated for health but does not include provision of water and sanitation.	World Bank World Development Indicators database (WDI 2014).
Neonatal mortality rate	Neonatal mortality rate is the number of neonates dying before reaching 28 days of age, per 1,000 live births in a given year.	WDI 2014
GDP per capita	GDP per capita, PPP (constant 2011 international $). PPP GDP is gross domestic product converted to international dollars using purchasing power parity rates.	WDI 2014
Age dependency ratio (of working age population)	Age dependency ratio is the ratio of dependants—people younger than 15 or older than 64—to the working age population—those aged 15–64. Data are shown as the proportion of dependants per 100 working age population.	WDI 2014
Military expenditure per GDP	Military expenditures data from SIPRI are derived from the NATO definition, which includes all current and capital expenditures on the armed forces, including peacekeeping forces; defence ministries and other government agencies engaged in defence projects; paramilitary forces, if these are judged to be trained and equipped for military operations; and military space activities.	WDI 2014
Population	Total population is based on the de facto definition of population, which counts all residents who are generally considered part of the population of their country of origin.	WDI 2014
Adolescent fertility	Adolescent fertility rate is the number of births per 1,000 women aged 15–19.	WDI 2014
Share of population under 15	Population between the ages 0 to 14 as a percentage of the total population. Population is based on the de facto definition of population.	WDI 2014
Female:male ratio in primary education	Gender parity index for gross enrolment ratio in primary education is the ratio of girls to boys enrolled at primary level in public and private schools.	WDI 2014

Source: World Development Indicators, the World Bank.

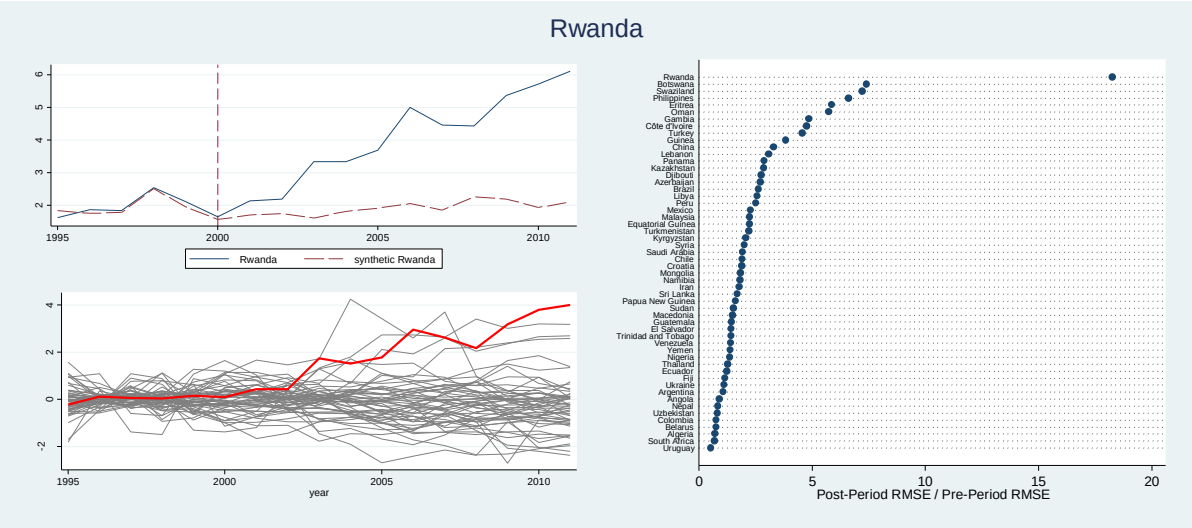
Figure A1: Synthetic control and robustness testing for Tanzania



Notes: The upper left-hand graph describes the outcome variables of the treated country and its synthetic controls. The lower left-hand graph describes the placebo testing results where the solid line describes the outcome difference between each treated country and its synthetic control and the dashed lines describe the outcome difference between each of the treated country's potential controls and their synthetic control in placebo experiments. The right-hand side graph shows the relationship between RMSPE pre- and post-treatment of the actual treatment country and the placebo treatment countries.

Source: Author's own calculations.

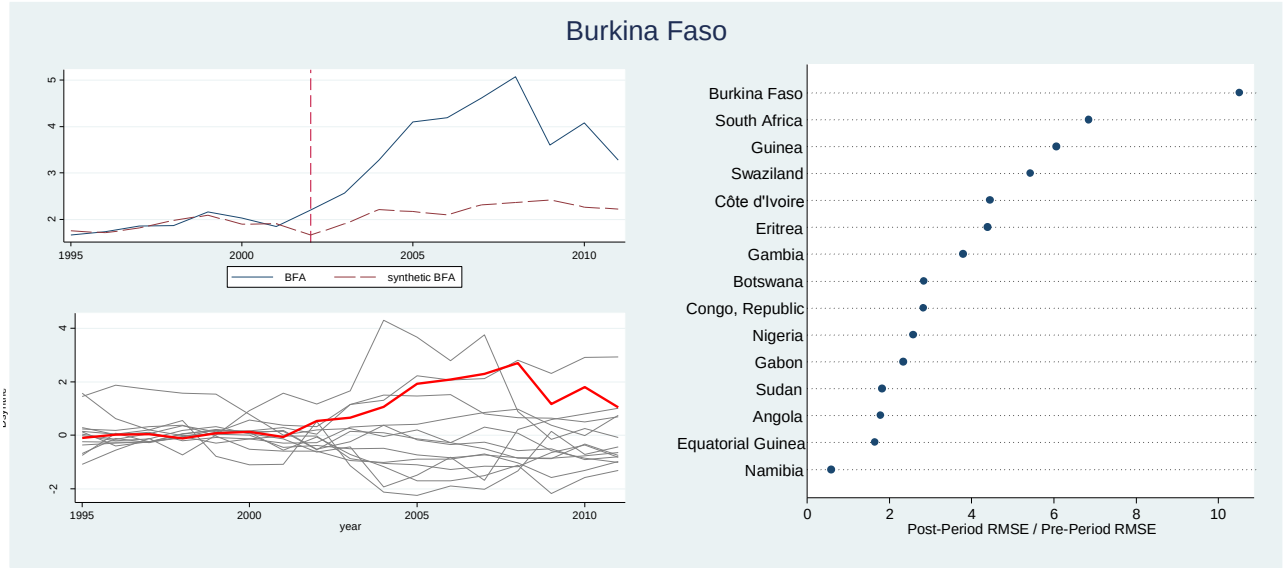
Figure A2: Synthetic control and robustness testing for Rwanda



Notes: The upper left-hand graph describes the outcome variables of the treated country and its synthetic controls. The lower left-hand graph describes the placebo testing results where the solid line describes the outcome difference between each treated country and its synthetic control and the dashed lines describe the outcome difference between each of the treated country's potential controls and their synthetic control in placebo experiments. The right-hand side graph shows the relationship between RMSPE pre- and post-treatment of the actual treatment country and the placebo treatment countries.

Source: Author's own calculations.

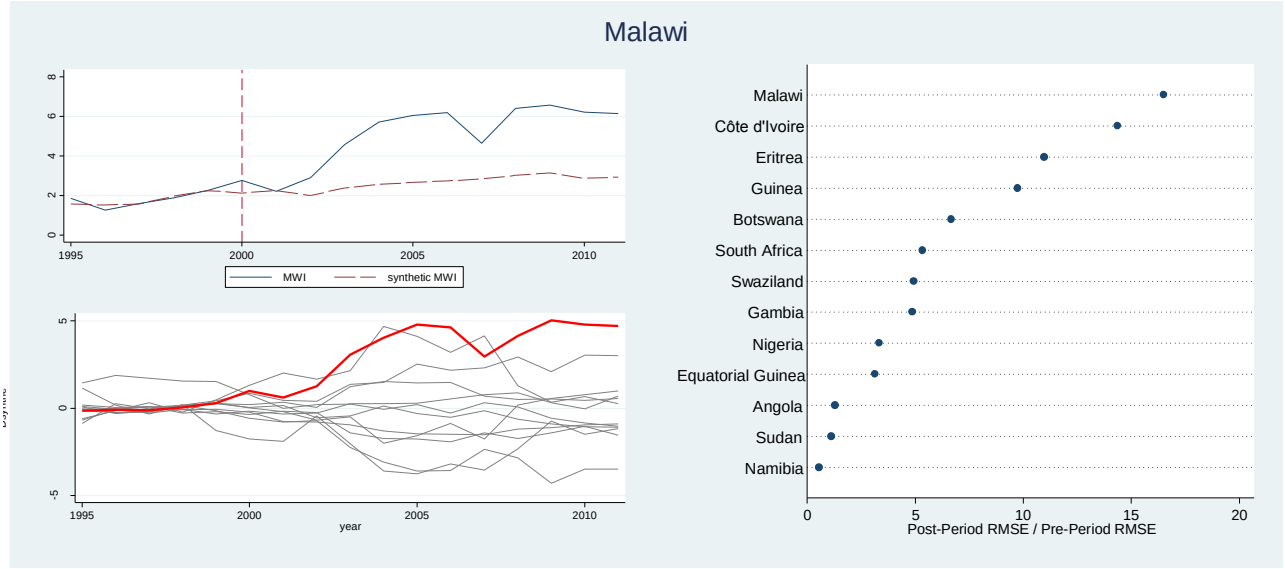
Figure A3: Synthetic control and robustness testing for Burkina Faso



Notes: The upper left-hand graph describes the outcome variables of the treated country and its synthetic controls. The lower left-hand graph describes the placebo testing results where the solid line describes the outcome difference between each treated country and its synthetic control and the dashed lines describe the outcome difference between each of the treated country's potential controls and their synthetic control in placebo experiments. The right-hand side graph shows the relationship between RMSPE pre- and post-treatment of the actual treatment country and the placebo treatment countries.

Source: Author's own calculations.

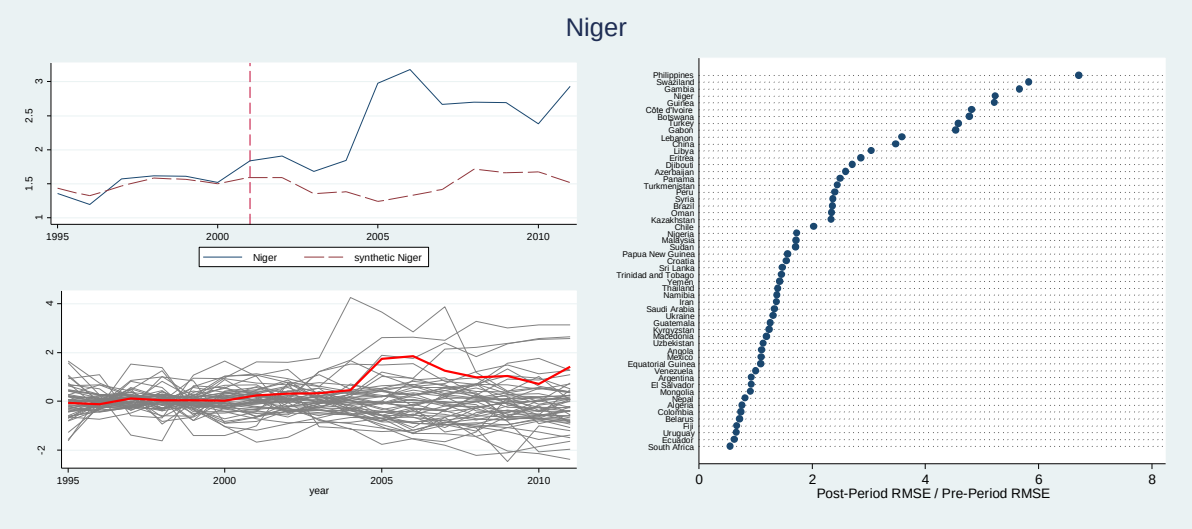
Figure A4: Synthetic control and robustness testing for Malawi



Notes: The upper left-hand graph describes the outcome variables of the treated country and its synthetic controls. The lower left-hand graph describes the placebo testing results where the solid line describes the outcome difference between each treated country and its synthetic control and the dashed lines describe the outcome difference between each of the treated country's potential controls and their synthetic control in placebo experiments. The right-hand side graph shows the relationship between RMSPE pre- and post-treatment of the actual treatment country and the placebo treatment countries.

Source: Author's own calculations.

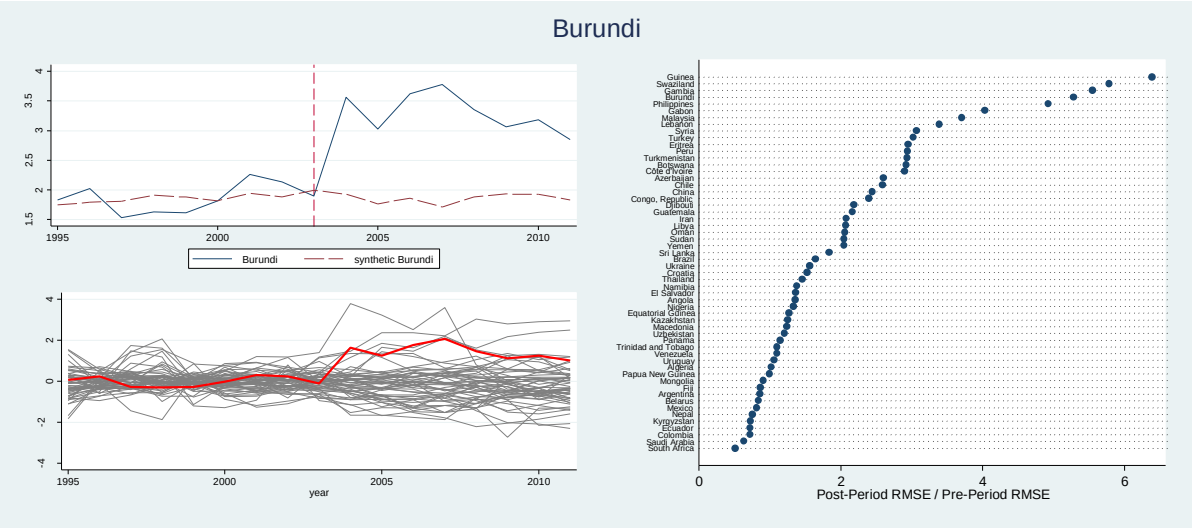
Figure A5: Synthetic control and robustness testing for Niger



Notes: The upper left-hand graph describes the outcome variables of the treated country and its synthetic controls. The lower left-hand graph describes the placebo testing results where the solid line describes the outcome difference between each treated country and its synthetic control and the dashed lines describe the outcome difference between each of the treated country's potential controls and their synthetic control in placebo experiments. The right-hand side graph shows the relationship between RMSPE pre- and post-treatment of the actual treatment country and the placebo treatment countries.

Source: Author's own calculations.

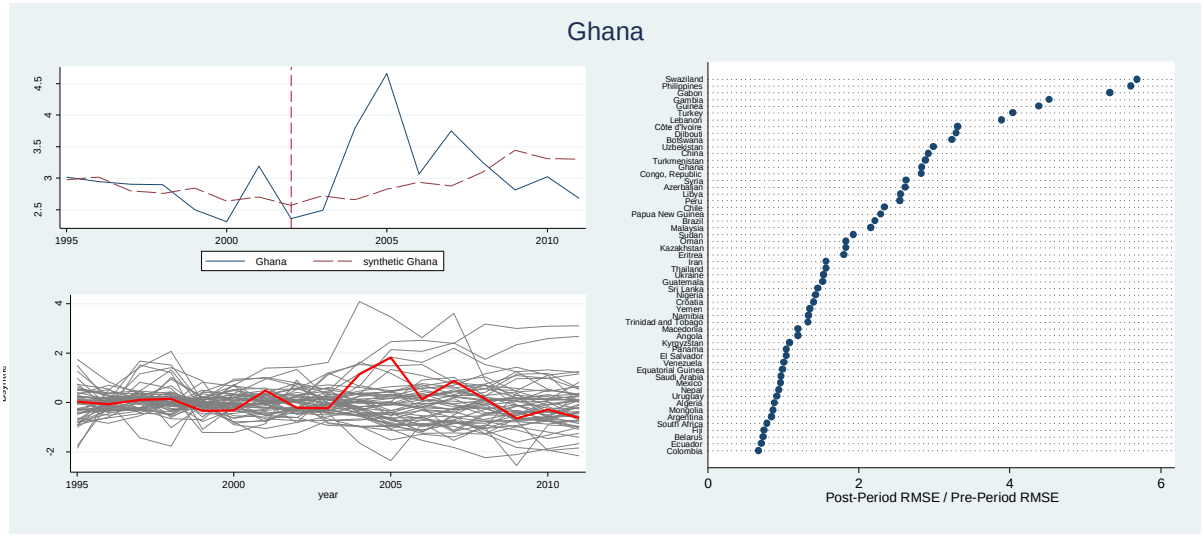
Figure A6: Synthetic control and robustness testing for Burundi



Notes: The upper left-hand graph describes the outcome variables of the treated country and its synthetic controls. The lower left-hand graph describes the placebo testing results where the solid line describes the outcome difference between each treated country and its synthetic control and the dashed lines describe the outcome difference between each of the treated country's potential controls and their synthetic control in placebo experiments. The right-hand side graph shows the relationship between RMSPE pre- and post-treatment of the actual treatment country and the placebo treatment countries.

Source: Author's own calculations.

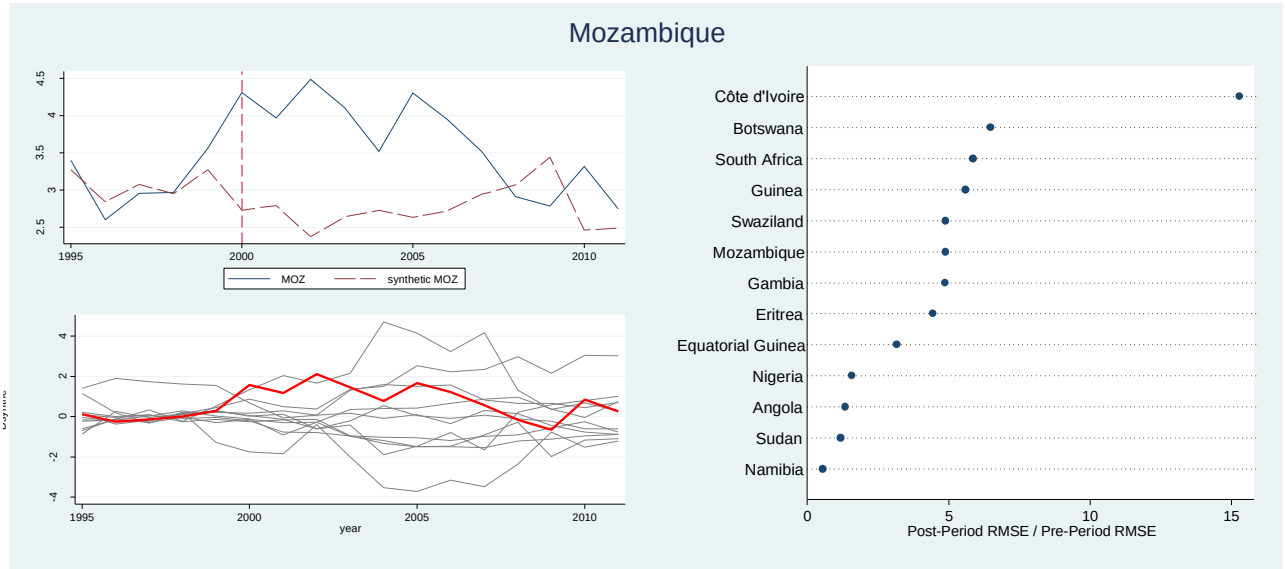
Figure A7: Synthetic control and robustness testing for Ghana



Notes: The upper left-hand graph describes the outcome variables of the treated country and its synthetic controls. The lower left-hand graph describes the placebo testing results where the solid line describes the outcome difference between each treated country and its synthetic control and the dashed lines describe the outcome difference between each of the treated country's potential controls and their synthetic control in placebo experiments. The right-hand side graph shows the relationship between RMSPE pre- and post-treatment of the actual treatment country and the placebo treatment countries.

Source: Author's own calculations.

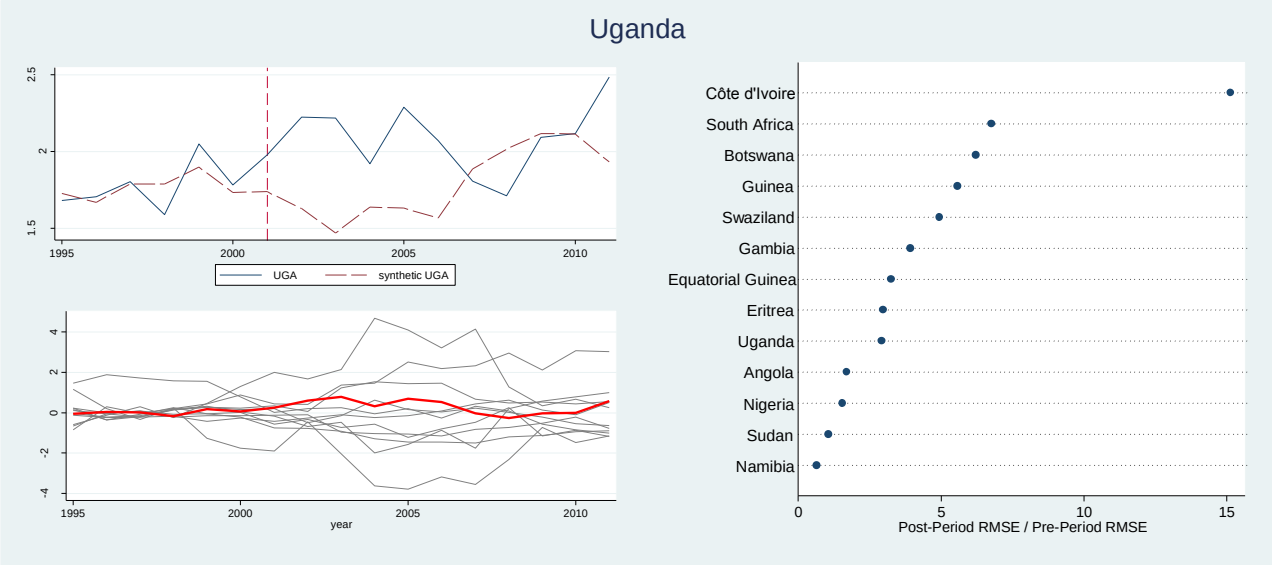
Figure A8: Synthetic control and robustness testing for Mozambique



Notes: The upper left-hand graph describes the outcome variables of the treated country and its synthetic controls. The lower left-hand graph describes the placebo testing results where the solid line describes the outcome difference between each treated country and its synthetic control and the dashed lines describe the outcome difference between each of the treated country's potential controls and their synthetic control in placebo experiments. The right-hand side graph shows the relationship between RMSPE pre- and post-treatment of the actual treatment country and the placebo treatment countries.

Source: Author's own calculations.

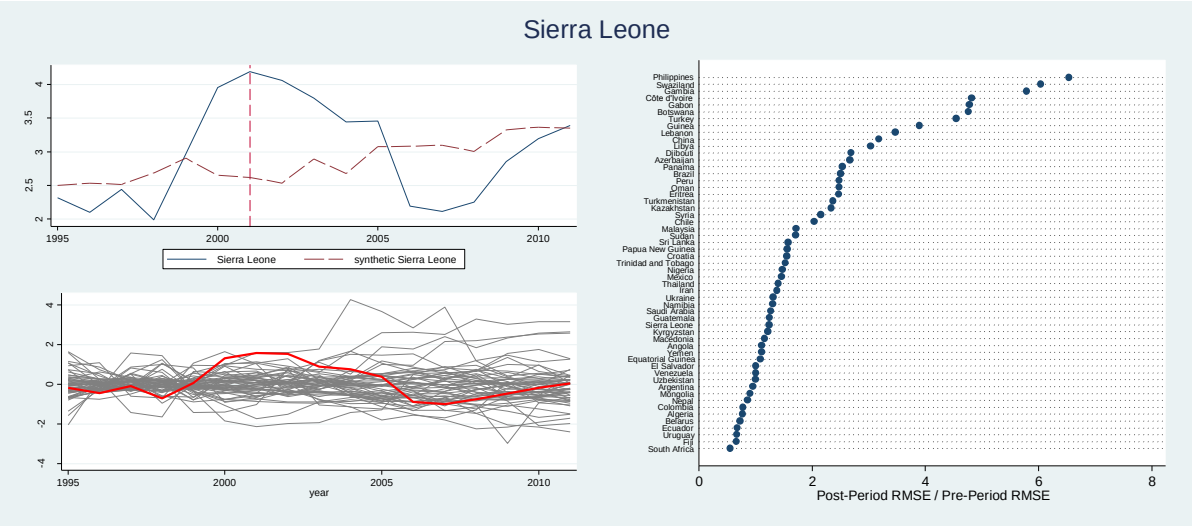
Figure A9: Synthetic control and robustness testing for Uganda



Notes: The upper left-hand graph describes the outcome variables of the treated country and its synthetic controls. The lower left-hand graph describes the placebo testing results where the solid line describes the outcome difference between each treated country and its synthetic control and the dashed lines describe the outcome difference between each of the treated country's potential controls and their synthetic control in placebo experiments. The right-hand side graph shows the relationship between RMSPE pre- and post-treatment of the actual treatment country and the placebo treatment countries.

Source: Author's own calculations.

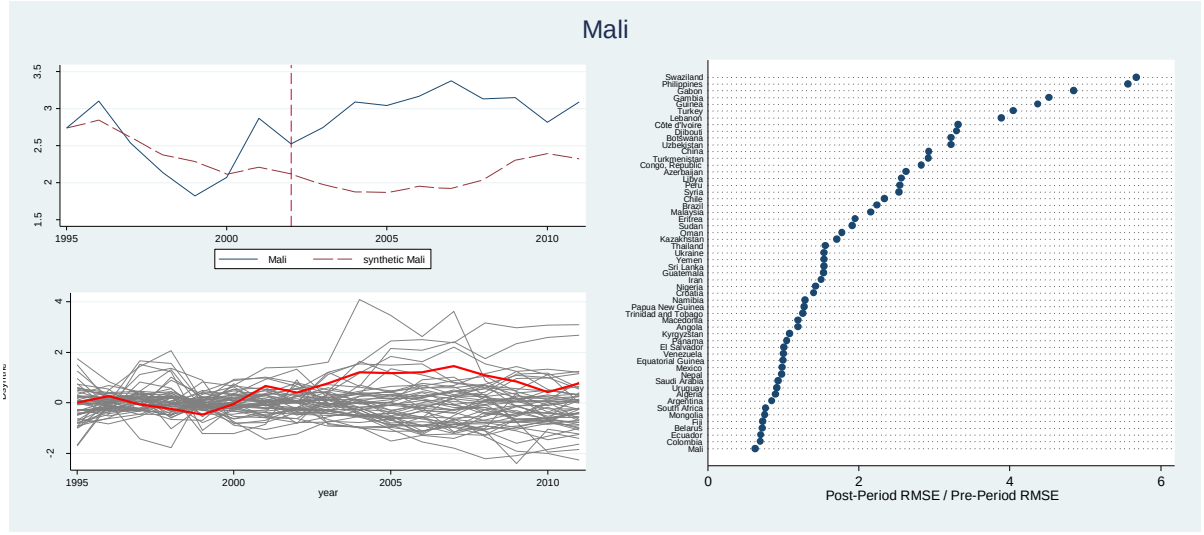
Figure A10: Synthetic control and robustness testing for Sierra Leone



Notes: The upper left-hand graph describes the outcome variables of the treated country and its synthetic controls. The lower left-hand graph describes the placebo testing results where the solid line describes the outcome difference between each treated country and its synthetic control and the dashed lines describe the outcome difference between each of the treated country's potential controls and their synthetic control in placebo experiments. The right-hand side graph shows the relationship between RMSPE pre- and post-treatment of the actual treatment country and the placebo treatment countries..

Source: Author's own calculations.

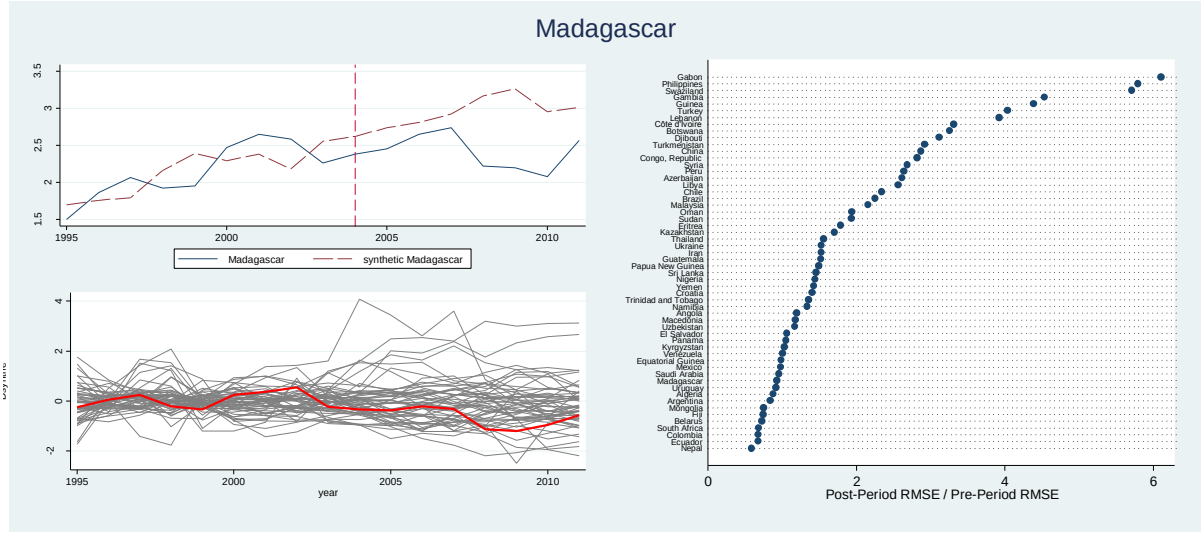
Figure A11: Synthetic control and robustness testing for Mali



Notes: The upper left-hand graph describes the outcome variables of the treated country and its synthetic controls. The lower left-hand graph describes the placebo testing results where the solid line describes the outcome difference between each treated country and its synthetic control and the dashed lines describe the outcome difference between each of the treated country's potential controls and their synthetic control in placebo experiments. The right-hand side graph shows the relationship between RMSPE pre- and post-treatment of the actual treatment country and the placebo treatment countries..

Source: Author's own calculations.

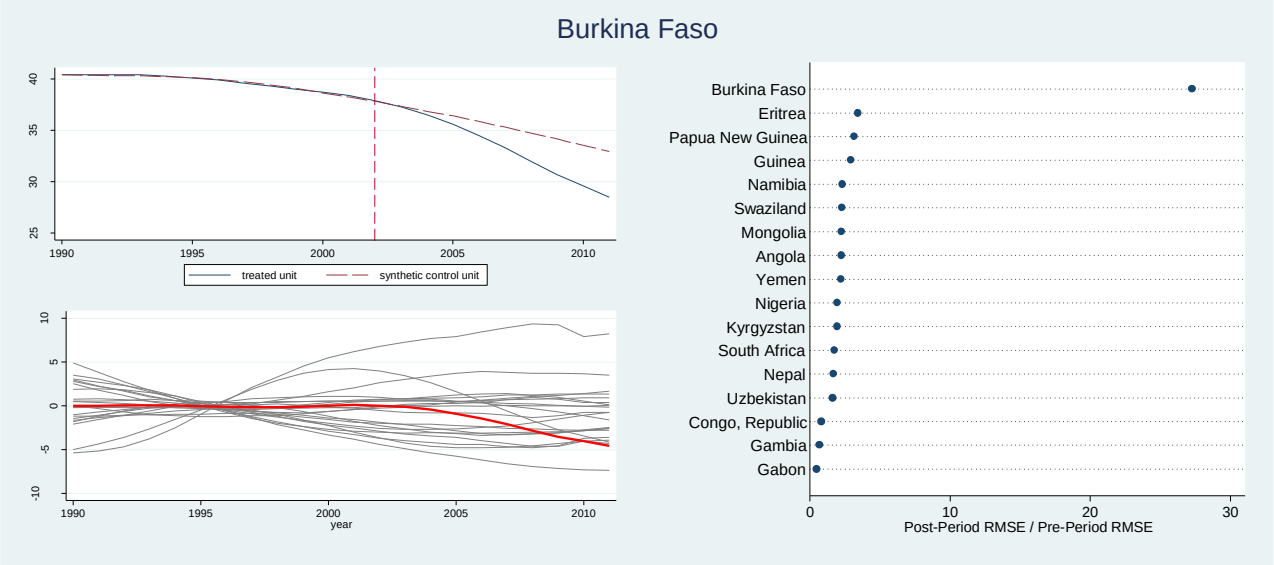
Figure A12: Synthetic control and robustness testing for Madagascar



Notes: The upper left-hand graph describes the outcome variables of the treated country and its synthetic controls. The lower left-hand graph describes the placebo testing results where the solid line describes the outcome difference between each treated country and its synthetic control and the dashed lines describe the outcome difference between each of the treated country's potential controls and their synthetic control in placebo experiments. The right-hand side graph shows the relationship between RMSPE pre- and post-treatment of the actual treatment country and the placebo treatment countries..

Source: Author's own calculations.

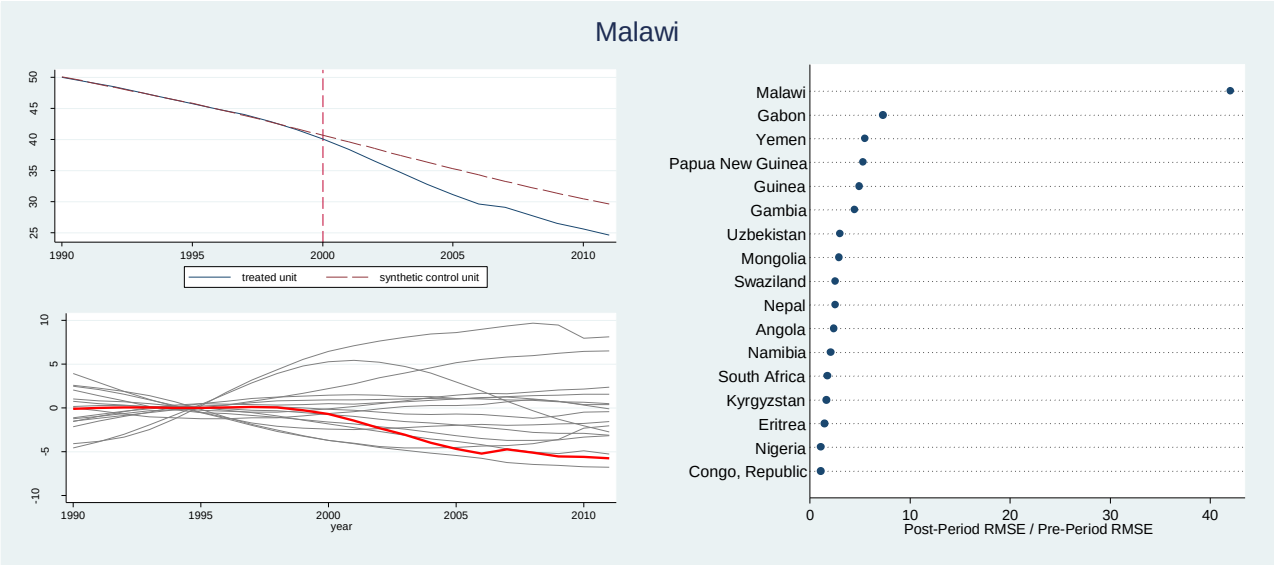
Figure A13: Burkina Faso: Neonatal mortality rate, synthetic control and robustness testing



Notes: The upper left-hand graph describes the outcome variables of the treated country and its synthetic controls. The lower left-hand graph describes the placebo testing results where the solid line describes the outcome difference between each treated country and its synthetic control and the dashed lines describe the outcome difference between each of the treated country's potential controls and their synthetic control in placebo experiments. The right-hand side graph shows the relationship between RMSPE pre- and post-treatment of the actual treatment country and the placebo treatment countries.

Source: Author's own calculations.

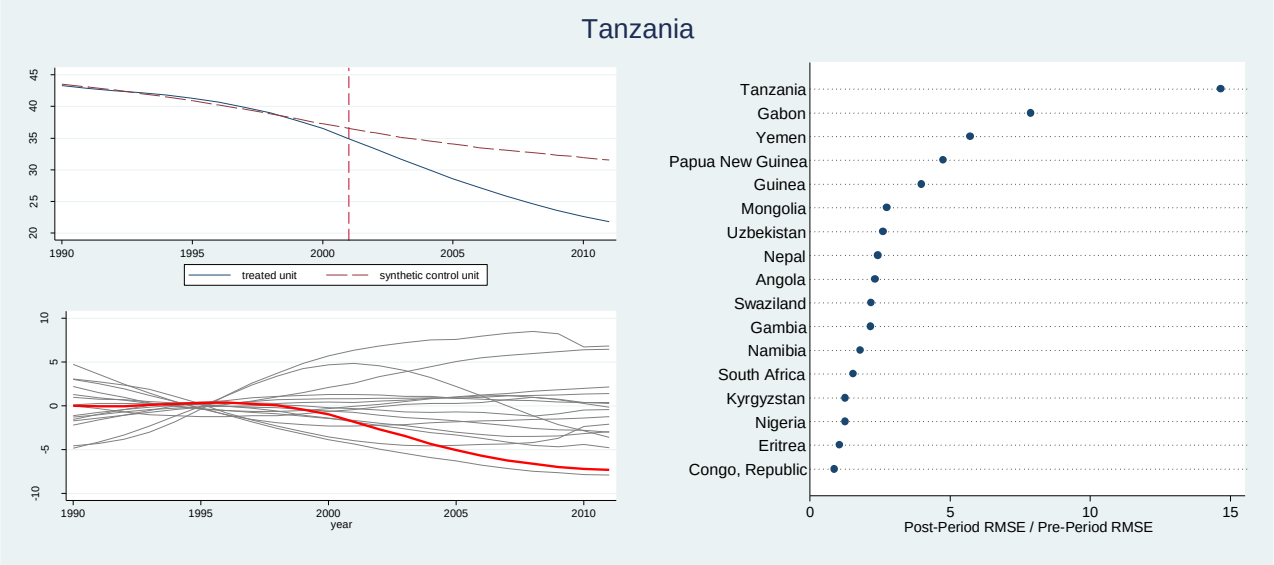
Figure A14: Malawi: Neonatal mortality rate, synthetic control and robustness testing



Notes: The upper left-hand graph describes the outcome variables of the treated country and its synthetic controls. The lower left-hand graph describes the placebo testing results where the solid line describes the outcome difference between each treated country and its synthetic control and the dashed lines describe the outcome difference between each of the treated country's potential controls and their synthetic control in placebo experiments. The right-hand side graph shows the relationship between RMSPE pre- and post-treatment of the actual treatment country and the placebo treatment countries.

Source: Author's own calculations.

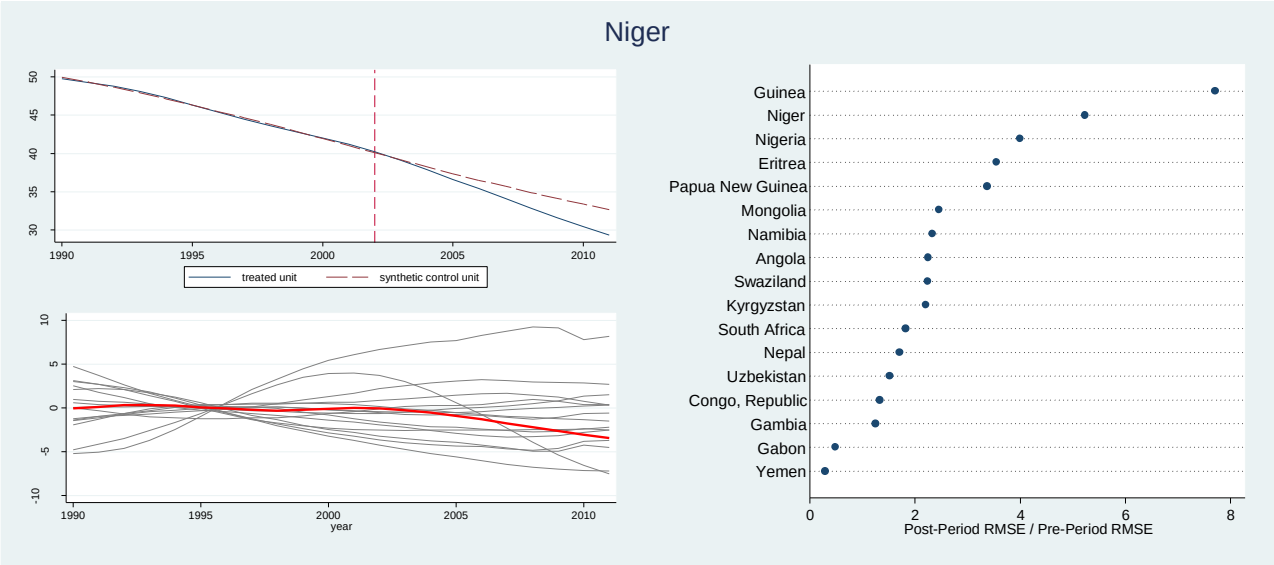
Figure A15: Tanzania: Neonatal mortality rate, synthetic control and robustness testing



Notes: The upper left-hand graph describes the outcome variables of the treated country and its synthetic controls. The lower left-hand graph describes the placebo testing results where the solid line describes the outcome difference between each treated country and its synthetic control and the dashed lines describe the outcome difference between each of the treated country's potential controls and their synthetic control in placebo experiments. The right-hand side graph shows the relationship between RMSPE pre- and post-treatment of the actual treatment country and the placebo treatment countries.

Source: Author's own calculations.

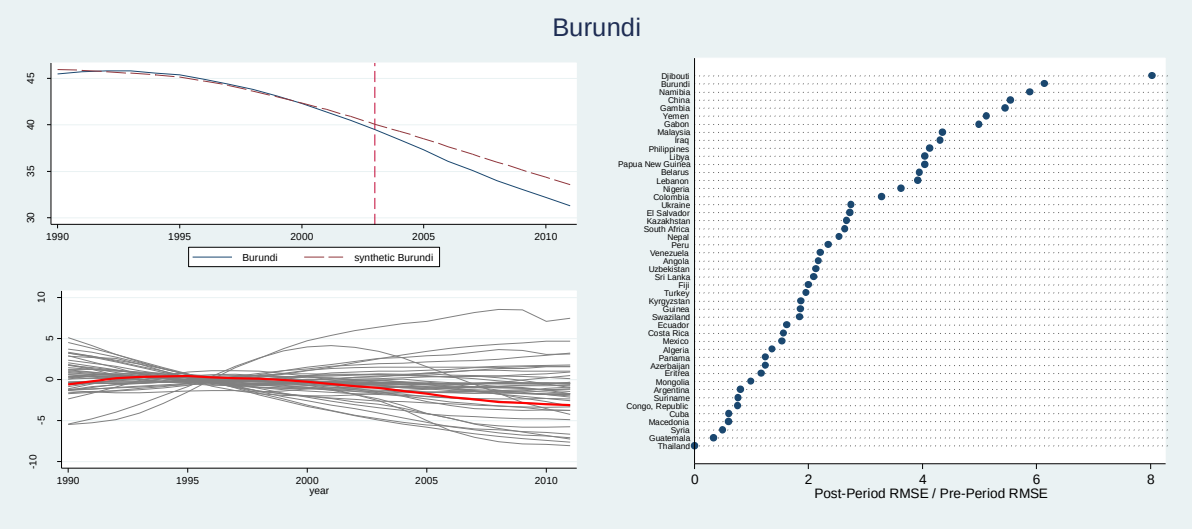
Figure A16: Niger: Neonatal mortality rate, synthetic control and robustness testing



Notes: The upper left-hand graph describes the outcome variables of the treated country and its synthetic controls. The lower left-hand graph describes the placebo testing results where the solid line describes the outcome difference between each treated country and its synthetic control and the dashed lines describe the outcome difference between each of the treated country's potential controls and their synthetic control in placebo experiments. The right-hand side graph shows the relationship between RMSPE pre- and post-treatment of the actual treatment country and the placebo treatment countries.

Source: Author's own calculations.

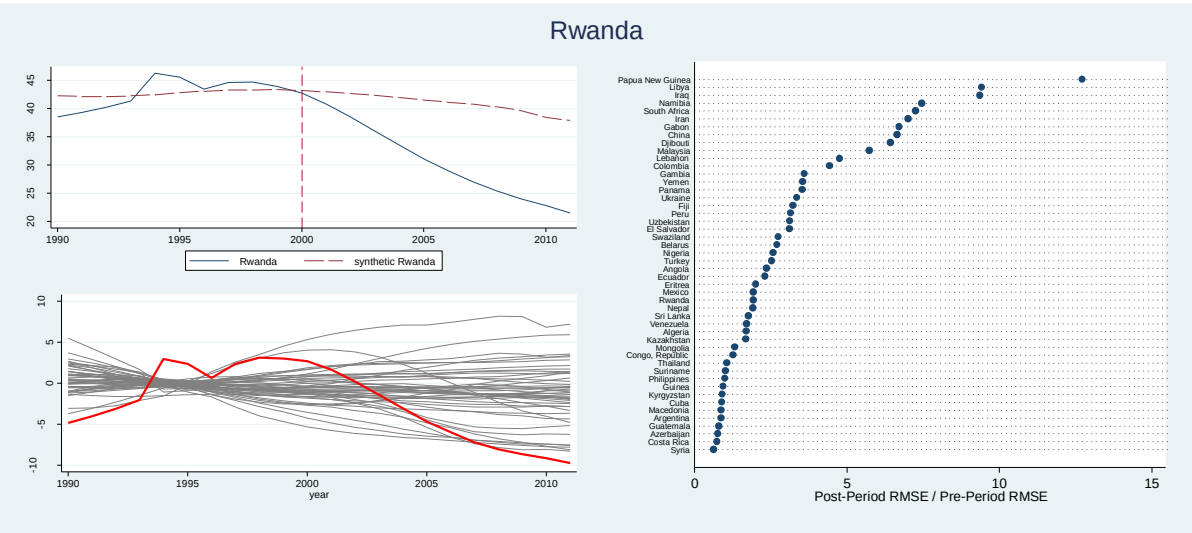
Figure A17: Burundi: Neonatal mortality rate, synthetic control and robustness testing



Notes: The upper left-hand graph describes the outcome variables of the treated country and its synthetic controls. The lower left-hand graph describes the placebo testing results where the solid line describes the outcome difference between each treated country and its synthetic control and the dashed lines describe the outcome difference between each of the treated country's potential controls and their synthetic control in placebo experiments. The right-hand side graph shows the relationship between RMSPE pre- and post-treatment of the actual treatment country and the placebo treatment countries.

Source: Author's own calculations.

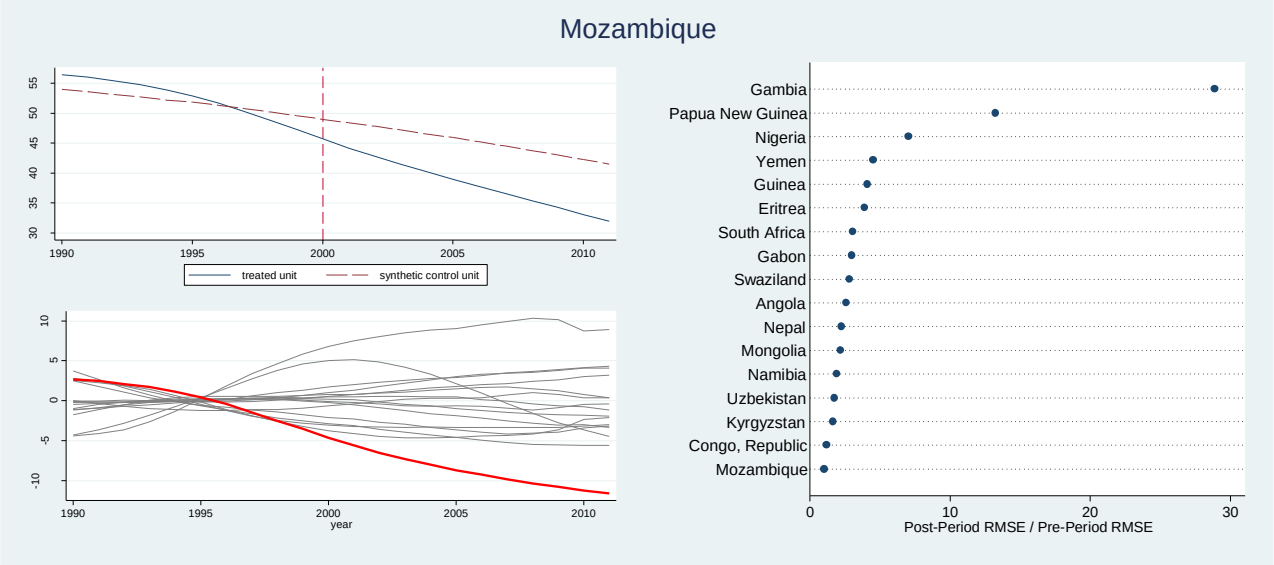
Figure A18: Rwanda: Neonatal mortality rate, synthetic control and robustness testing



Notes: The upper left-hand graph describes the outcome variables of the treated country and its synthetic controls. The lower left-hand graph describes the placebo testing results where the solid line describes the outcome difference between each treated country and its synthetic control and the dashed lines describe the outcome difference between each of the treated country's potential controls and their synthetic control in placebo experiments. The right-hand side graph shows the relationship between RMSPE pre- and post-treatment of the actual treatment country and the placebo treatment countries.

Source: Author's own calculations.

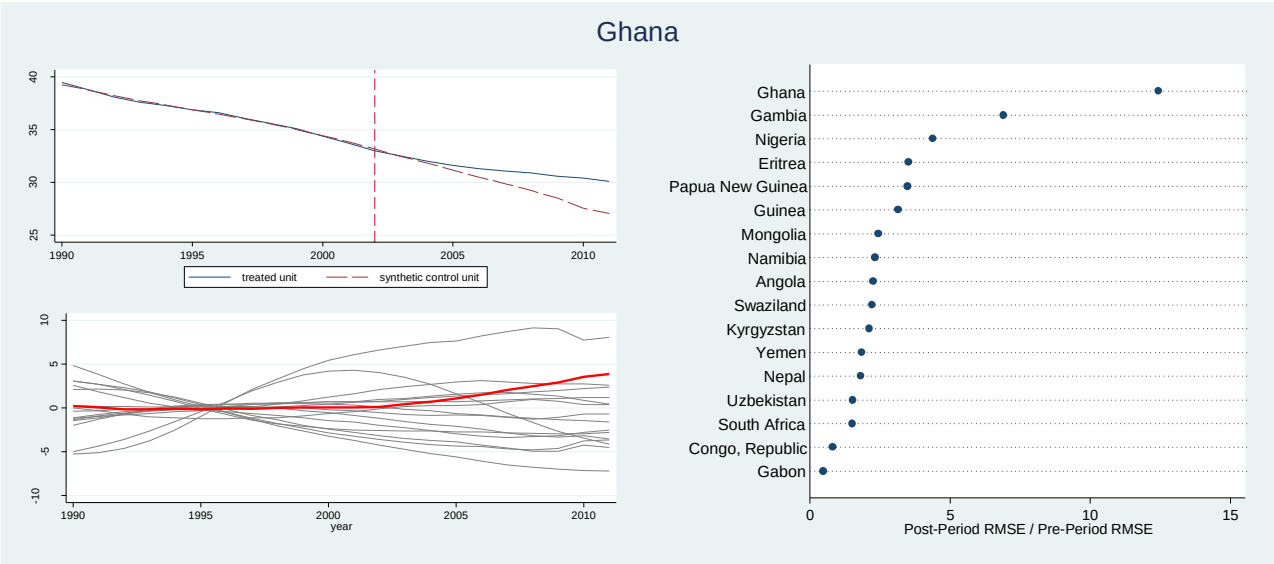
Figure A19: Mozambique: Neonatal mortality rate, synthetic control and robustness testing



Notes: The upper left-hand graph describes the outcome variables of the treated country and its synthetic controls. The lower left-hand graph describes the placebo testing results where the solid line describes the outcome difference between each treated country and its synthetic control and the dashed lines describe the outcome difference between each of the treated country's potential controls and their synthetic control in placebo experiments. The right-hand side graph shows the relationship between RMSPE pre- and post-treatment of the actual treatment country and the placebo treatment countries.

Source: Author's own calculations.

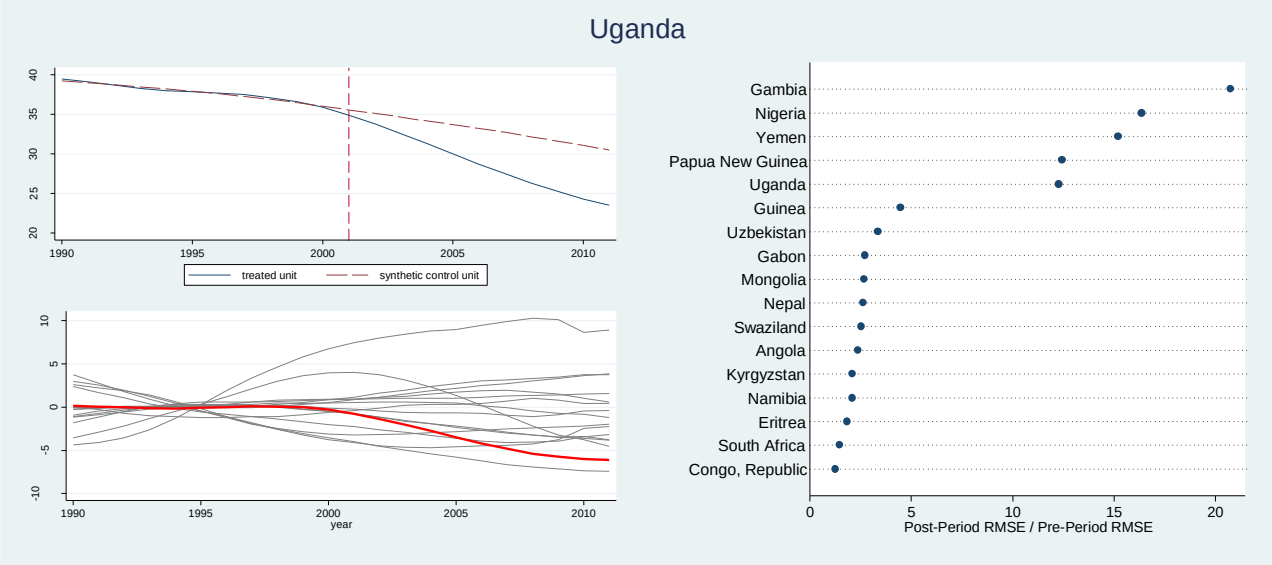
Figure A20: Ghana: Neonatal mortality rate, synthetic control and robustness testing



Notes: The upper left-hand graph describes the outcome variables of the treated country and its synthetic controls. The lower left-hand graph describes the placebo testing results where the solid line describes the outcome difference between each treated country and its synthetic control and the dashed lines describe the outcome difference between each of the treated country's potential controls and their synthetic control in placebo experiments. The right-hand side graph shows the relationship between RMSPE pre- and post-treatment of the actual treatment country and the placebo treatment countries.

Source: Author's own calculations.

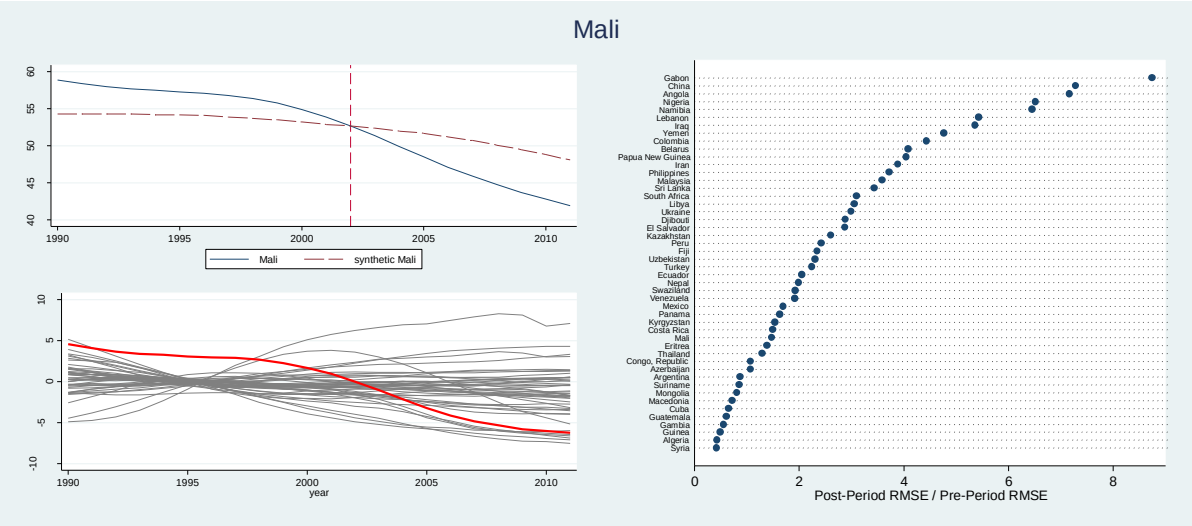
Figure A21: Uganda: Mortality rate, synthetic control and robustness testing



Notes: The upper left-hand graph describes the outcome variables of the treated country and its synthetic controls. The lower left-hand graph describes the placebo testing results where the solid line describes the outcome difference between each treated country and its synthetic control and the dashed lines describe the outcome difference between each of the treated country's potential controls and their synthetic control in placebo experiments. The right-hand side graph shows the relationship between RMSPE pre- and post-treatment of the actual treatment country and the placebo treatment countries.

Source: Author's own calculations.

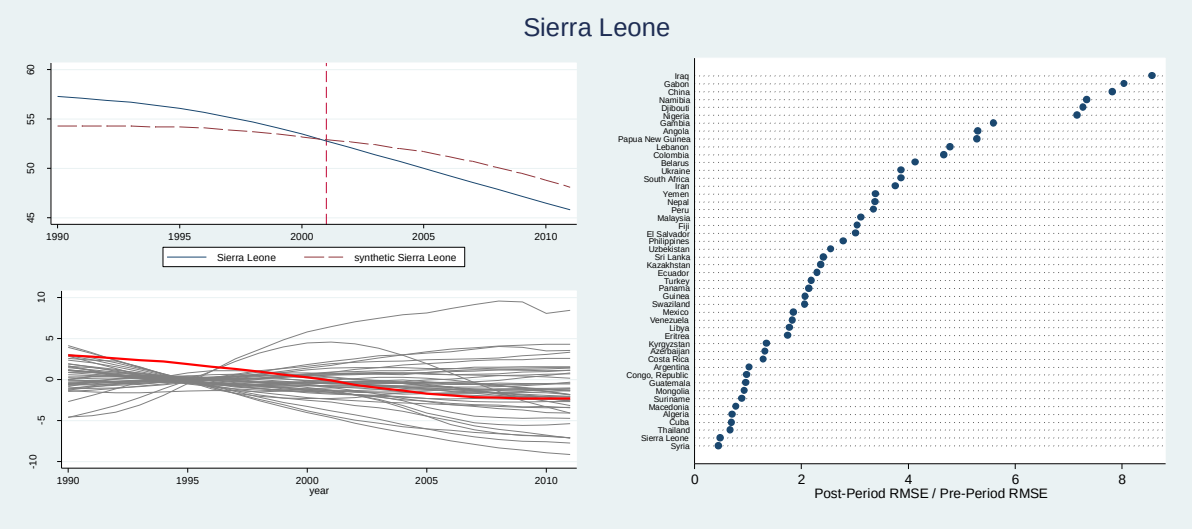
Figure A22: Mali: Mortality rate, synthetic control and robustness testing



Notes: The upper left-hand graph describes the outcome variables of the treated country and its synthetic controls. The lower left-hand graph describes the placebo testing results where the solid line describes the outcome difference between each treated country and its synthetic control and the dashed lines describe the outcome difference between each of the treated country's potential controls and their synthetic control in placebo experiments. The right-hand side graph shows the relationship between RMSPE pre- and post-treatment of the actual treatment country and the placebo treatment countries.

Source: Author's own calculations.

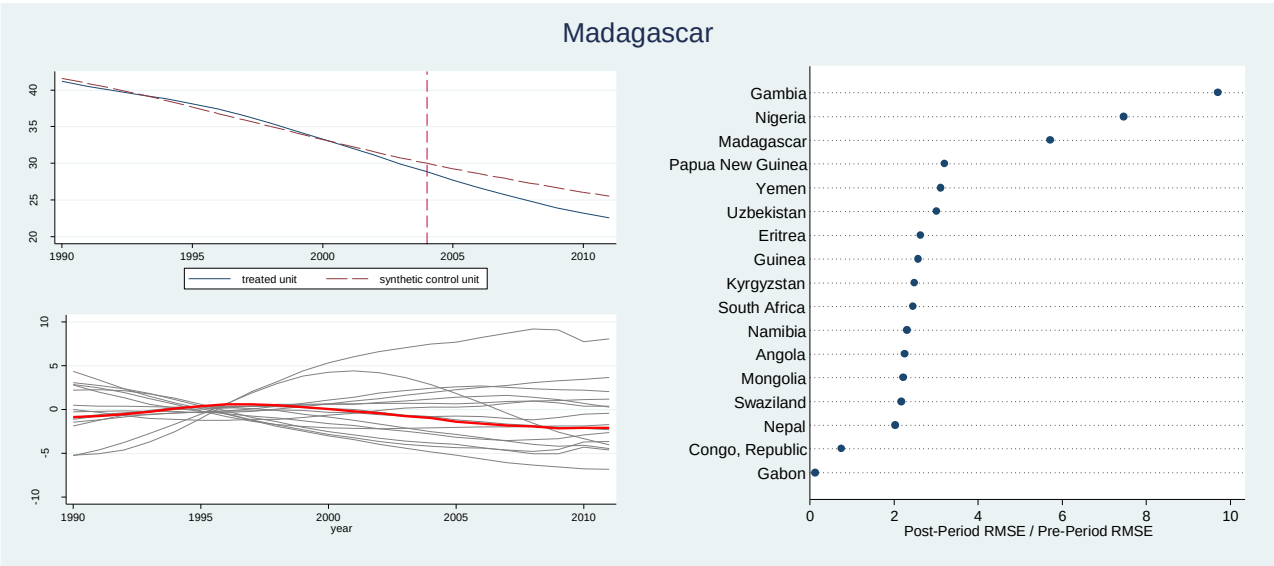
Figure A23: Sierra Leone: Neonatal mortality rate, synthetic control and robustness testing



Notes: The upper left-hand graph describes the outcome variables of the treated country and its synthetic controls. The lower left-hand graph describes the placebo testing results where the solid line describes the outcome difference between each treated country and its synthetic control and the dashed lines describe the outcome difference between each of the treated country's potential controls and their synthetic control in placebo experiments. The right-hand side graph shows the relationship between RMSPE pre- and post-treatment of the actual treatment country and the placebo treatment countries.

Source: Author's own calculations.

Figure A24: Sierra Leone: Neonatal mortality rate, synthetic control and robustness testing



Notes: The upper left-hand graph describes the outcome variables of the treated country and its synthetic controls. The lower left-hand graph describes the placebo testing results where the solid line describes the outcome difference between each treated country and its synthetic control and the dashed lines describe the outcome difference between each of the treated country's potential controls and their synthetic control in placebo experiments. The right-hand side graph shows the relationship between RMSPE pre- and post-treatment of the actual treatment country and the placebo treatment countries.

Source: Author's own calculations.