STOCKHOLM SCHOOL OF ECONOMICS
Department of Economics
659 Degree project in economics
Spring 2022

The impact of ICT diffusion on primary schooling in The Gambia: a synthetic control approach

Ida Lennström (24309) and Spencer Robild (24485)

Abstract: Using national-level data from countries within sub-Saharan Africa, this paper investigates the impact of ICT diffusion, measured through Internet penetration, on primary school completion. Digital inequalities, as well as an overrepresentation of children out of school, particularly affect poor- and rural population groups in sub-Saharan Africa. While ICT has been identified as essential in order to mitigate educational challenges, research on the relationship between ICT and education is lacking. This study utilises the Synthetic Control Method in combination with a critical threshold of Internet penetration to study the causality and aims to provide value for policymakers in relation to ICT deployment and developments of the education system. The evidence for the actual outcome of The Gambia compared with the synthetic counterfactual displays a negative treatment effect on primary school completion, albeit placebo- and robustness tests show weak validity. Limitations put forth by the SCM in the specific context yield a null result and leave room for future research to address the relationship between ICT and education. The authors encourage further research using the newly developed Augmented SCM, as well as broader research into the interplay of determinants of educational performance and educational attainment. Future studies should furthermore expand upon secondary and tertiary education, not covered in this study.

Keywords: Internet penetration, Synthetic Control Method, sub-Saharan Africa, primary school completion, ICT diffusion

JEL: I24, I25, O15, O33

Supervisor: Céline Zipfel
Date submitted: 16 May 2022
Date examined: 24 May 2022
Discussants: Olof de Blanche and Sofia Ferenius
Examiner: Johanna Wallenius

Acknowledgements

We would like to extend our deepest gratitude to our supervisor Céline Zipfel for invaluable support during the project. Furthermore, we would like to thank Marion Leroutier and Valter Frankenberg for helpful feedback regarding the empirical methodology. Lastly, we are grateful for critical comments from our discussants and classmates.

Concepts and definitions

GDP = Gross Domestic Product

ICT = Information and Communication Technology

SDG = Sustainable Development Goals, adopted by the United Nations

GA = The Gambia

SSA = Sub-Saharan Africa

SGA = Synthetic The Gambia

AC = Average Control

LOO = Leave-One-Out

SCM = Synthetic Control Method

RMSPE = Root Mean Square Prediction Error

Contents

List of Figures

List of Tables

1 Introduction

As an integral part of developed societies today, the Internet has spread quicker than any other information and communication technology (ICT) globally. The "digital divide" has surfaced as a term for the digital inequalities present between- and within countries with only 29 per cent of the population in sub-Saharan Africa (SSA) using the Internet in 2019 (see Figure 1; Rainie and Wellman 2019, 27–38; Roser, Ritchie, and Ortiz-Ospina 2015). Furthermore, the population group most affected by the digital divide is people with lower socioeconomic status (Rainie and Wellman 2019). Apart from pure information dissemination, ICT provides means to improve upon educational attainment and has been identified by the United Nations as essential in order to achieve the UN Sustainable Development Goal 4 (SDG 4). SDG 4 primarily focuses on ensuring that all children have free access to primary and secondary education (Lim et al. 2020; United Nations n.d.).

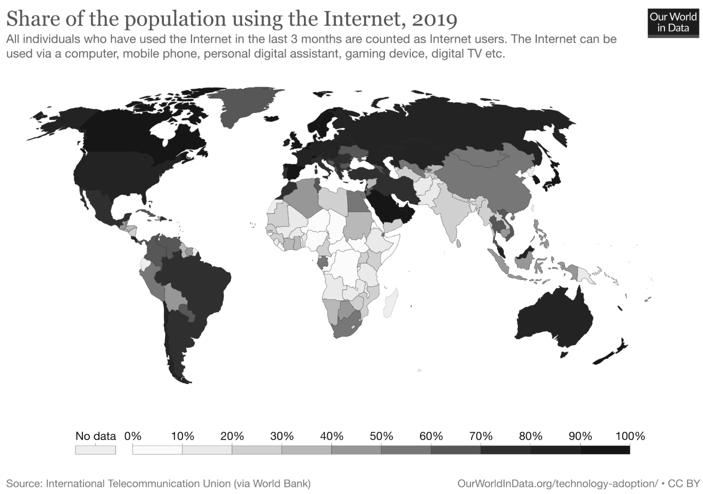

Figure 1: Share of the population using the Internet, 2019

Source: Roser, Ritchie, and Ortiz-Ospina 2015

Apart from the digital divide and the overrepresentation of children out of school in SSA, the region is affected by poor living standards, measured as the lowest GDP per capita out of all world regions (Roser 2013; Roser and Ortiz-Ospina 2013). According to UNESCO (2021), ICT can be used in the educational setting to disseminate, create, and store information. Through the use of digital whiteboards, smartphones and the flipped classroom model, where students listen to lectures at home and have interactive classes in school, ICT can be an essential part of the teaching experience (UNESCO 2021). Furthermore, some SSA countries struggle with declining access to textbooks, and e-readers can be used to store hundreds of books and deliver course material to students (UNESCO 2021; 'The Gambia Education Sector Strategic Plan 2016-2030' 2017). In addition, interventions where teachers are connected with children in rural areas through the use of ICT have yielded positive results (Bianchi, Song, and Lu 2020). Several mechanisms through which ICT have the potential to impact education thus highlight the need for comprehensive research on the topic.

While many scholars have studied causal relations between economic growth and ICT development, as well as between economic growth and educational attainment, less research has been conducted on the relationship between ICT diffusion and education. Has the notable digital divide in SSA had negative impacts on educational completion for children? Such questions are of interest for policy applications in developing countries and would be of value to decision-makers in regard to ICT development. This paper aims to study the effects of Internet penetration on primary schooling in SSA.

In recent years, the Synthetic Control Method (SCM) has gained popularity due to its advantages in statistically deriving the causal effect in comparative case studies (Abadie and Gardeazabal 2003; Abadie, Diamond, and Hainmueller 2010; Abadie 2021). In our paper, the causal effect of Internet penetration on primary school completion rate can be studied by utilising the SCM. Since data availability restricts us to national-level panel data provided by the World Bank, the SSA, with varying degrees of national Internet penetration, provides a favourable sample for a comparative case study. One of the countries with above-average Internet penetration is the Republic of The Gambia (The Gambia), with 35 per cent Internet penetration[1] in 2018 compared to the average of 25 per cent for SSA (World Bank n.d.a). Our study draws upon existing research on the critical threshold of Internet penetration[2], to practically apply the SCM in a case where intervention occurs gradually, such as the case with technology adoption.

The SCM provides a method for the creation of an artificial counterfactual of The Gambia. Synthetic The Gambia (SGA) will consist of a weighted combination of sub-Saharan African countries lacking above-threshold Internet penetration, but that previously have resembled The Gambia in terms of different key predictors. By comparing the primary school completion rate in The Gambia (GA) with the synthetic counterfactual, the effect of above-threshold Internet penetration on the primary school completion rate can be assessed and studied.

[1] Measured as the percentage of the population using the Internet.
[2] The critical mass of Internet users needed to ensure exponential network growth as well as continued positive network externalities further explained in section 4.4.

Our findings display a largely negative treatment effect post-intervention, meaning that above-threshold Internet penetration negatively affects primary completion. Placebo tests together with a robustness test show weak validity for the treatment effect, however, subsequently meaning that a null result is yielded, rejecting our hypothesis of a positive relationship. Previous literature presents contrasting empirical evidence for the relationship between ICT and education[3], which is why the null result is not considered unexpected. Other determinants seem to greatly impact primary completion, and being able to isolate the causal relationship in this setting thus proves concerning. The authors propose further research related to different determinants and their interplay, as well as future research expanded to differentiate between levels of education. A newly developed addition to the SCM, the Augmented SCM (cf. Ben-Michael, Feller, and Rothstein 2021), could moreover provide value to the empirical setting.

Since no previous studies have applied the SCM to empirically study the relationship between ICT and educational attainment[4], our findings add to the existing research through the use of a new methodology. Furthermore, this paper contributes to the broader topic of ICT usage in the educational setting, providing great value for policymakers in relation to ICT deployment and developments of the education system.

The remaining sections of this paper will be organised in the following way. Section 2 provides an overview of The Gambia, the macro-region of SSA, historical ICT developments and educational challenges. Section 3 provides a literature review, summarising previous research on the topic of ICT in relation to educational outcomes. Section 4 stipulates the research question, our hypotheses and the empirical methodology using the SCM. Section 5 outlays the datasets used in the study. Section 6 presents the empirical results in three different parts: the construction of counterfactual The Gambia, the effect on primary school completion rate, and placebo- and robustness tests. Section 7 provides a discussion of the results presented as well as future research suggestions, comparing and contrasting the findings with previous studies. Lastly, Section 8 outlays our final remarks.

[3] See section 3.2.
[4] To the best of the authors' knowledge at the time of writing this paper.

2 Background

In order to get an accurate understanding of the relationship between ICT and education in SSA, an understanding of the macro-region itself and the development of ICT and the educational system is necessary. Firstly, an overview of SSA and The Gambia specifically will be outlaid. Secondly, ICT development in SSA will be further examined, and lastly, educational reforms and educational outcomes will be studied.

2.1 Sub-Saharan Africa and The Gambia

SSA is defined as one of the world's macro-regions, delimited by factors of economic development, geopolitics, history, culture and sociology. While most scholars agree that the continent of Africa consists of two macro-regions, namely sub-Saharan Africa and the Middle East[5], the delimitation of countries included in SSA differs to some degree (Anděl, Bičík, and Bláha 2018). In this study, the World Bank's delimitation of countries included in SSA is used since datasets are derived from the same source (World Bank n.d.c).

As of 2020, SSA is the poorest region globally (see Figure 2), with little possibility of eradicating extreme poverty before 2030 (Bicaba, Brixiová, and Ncube 2017). The region has recently experienced democratisation, a large spread of new technologies, discoveries of natural resources and a decreased number of armed conflicts - all contributing to what could be considered a turning point for the region (UNDP n.d.).

Figure 2: Gross Domestic Product, 2020

Source: Roser 2013

[5] The Middle East; the Islamic Region; South-Western Asia and Northern Africa are all used interchangeably in the literature.

According to the World Bank, The Gambia is located in Western Africa and surrounded by Senegal and the Atlantic Ocean. With a population of approximately 2.5 million, the country is one of the most densely populated in the region. The agricultural labour market in The Gambia contributes to economic growth, but the poverty rate was 8.5 per cent in 2021 despite economic growth (World Bank n.d.d).

2.2 ICT developments in sub-Saharan Africa

Roser, Ritchie and Ortiz-Ospina (2015) express that the development of the World Wide Web in 1989 constituted a major milestone for global communication. By 2000, Norway and Canada had already reached an Internet penetration of 50 per cent, while SSA only had an Internet penetration of approximately one per cent. Globally, 640 000 people become first time Internet users on any given day on average, although nearly half of the global population is still offline. ICT diffusion[6] entails both Internet penetration, broadband access, mobile subscriptions and other measures of connectivity (Roser, Ritchie, and Ortiz-Ospina 2015). The digital divide exists both between- and within countries and particularly affects people in rural areas, low-income households and central-city areas, as well as people with lower educational attainment and people with disabilities (Rainie and Wellman 2019).

Up until 2005, less than five per cent of the population in The Gambia used the Internet (World Bank n.d.a). While the region of SSA still is very disconnected from online communication, several successful countries, including The Gambia, have managed to dramatically increase their Internet penetration recently (Roser, Ritchie, and Ortiz-Ospina 2015). In the National Development Plan for 2018-2021, adopted by the Government of The Gambia, one of the key performance indicators is to ensure that the share of the population using the Internet on a daily basis reaches 90 per cent (Government of The Gambia 2017).

2.3 Educational challenges in sub-Saharan Africa

Roser and Ortiz-Ospina (2013) state that primary education provides an essential foundation for literacy and basic mathematical skills, vital for one's ability to independently comprehend information. In 2019, more than 58 million children were not attending primary school on a global level, and more than half of them reside in SSA (Roser and Ortiz-Ospina 2013). While the number of children in primary education has grown substantially over the last couple of decades,

[6] According to the Oxford Dictionary, the act of spreading something widely in all directions; the fact of being spread in all directions.

many children are still out of attendance or not completing primary education in SSA (Lewin 2009). While many scholars argue for the connection between primary education and economic development, factors such as religious affiliation also affect educational attainment in SSA (Daun 2000). The Millenial Development Goals and the Sustainable Development Goals, both developed and adopted by the United Nations, have put emphasis on ensuring access to primary education for all children globally (United Nations n.d., 4).

According to the governmental Education Sector Strategic Plan for 2016-2030, The Gambia faces several challenges of low educational attainment, uneven access to schooling, low quality of education and declining access to school textbooks. Primary education starts in The Gambia at the age of seven, followed by six years of lower basic education and three years of subsequent upper basic education. Children in rural areas, as well as the poorest children in both urban and rural areas, are currently the groups where access to education is most lacking. Despite many educational improvements in recent decades, the average level of educational attainment within the population was 3.7 years in 2015. Due to declining textbook access, the responsible ministries for primary, secondary and tertiary education plan to introduce digital teaching with elements of ICT included in the curriculum during the ongoing policy period ('The Gambia Education Sector Strategic Plan 2016-2030' 2017). Since both educational and digital inequalities mostly affect low-income and rural population groups, the impact of ICT on education might be all the more important to study further.

3 Literature review

ICT diffusion, in this study measured by Internet penetration, is a widely studied topic by many scholars. Correlation studies associate ICT diffusion with ICT infrastructure, literacy, economic development and access fees. The interplay of different correlating factors, however, is what subsequently drives further ICT diffusion (A. Dutta and Roy 2005). ICT has spread exponentially over the past decades, contributing to economic, social and political development. The inequalities in regard to ICT availability, however, are one of the common explaining factors for the inaccessibility to healthcare services, education and commerce (Ayanso, Cho, and Lertwachara 2014).

Furthermore, A. Dutta and Roy (2005) outline the mechanics of ICT diffusion in their paper, arguing that an increase in ICT supply will not automatically lead to increased usage. Ayanso, Cho, and Lertwachara (2014) identify four different categories of access barriers with regard to the adoption of ICT. A *mental access barrier* can arise due to a lack of interest, a *material access barrier* can arise due to a lack of hardware and connectivity, a *skill access barrier* can arise due to a lack of skills education, and a *usage access barrier* can arise due to expensive usage fees among other factors. To address access barriers, A. Dutta and Roy (2005) suggest that countries need to adopt education- and training initiatives, demonstrate ICT capabilities through government services, increase geographical access and reduce access fees.

3.1 ICT diffusion positively contributes toward economic growth

The relationship between ICT diffusion and economic development has been subject to many studies other the last decades (cf. Stanley, Doucouliagos, and Steel 2018; Niebel 2014; David and Grobler 2020; Albiman and Sulong 2016; Hjort and Tian 2021; Baliamoune-Lutz 2003), all reaching a general consensus of an existing positive correlation between increased ICT diffusion and economic growth.

Stanley, Doucouliagos, and Steel (2018) provide findings for the positive contribution of ICT (specifically landline and cell technologies) to economic growth but emphasise that developed countries reap more benefits from ICT connectivity than developing countries. Adding to the debate, Baliamoune-Lutz (2003) provides empirical findings for a correlation between economic growth and ICT dissemination, stating that ICT contributes positively to the development and civil rights.

Baliamoune-Lutz (2003) argues that ICT could indeed impact economic growth positively, but that since the successful diffusion of ICT is determined by factors of social development and trade policies, developing countries might not succeed in the process of leapfrogging[7] through means of ICT. Niebel (2014) further questions the concept of leapfrogging through means of ICT, arguing that there is no empirical support for a significant difference between developed and developing countries.

Additionally, David and Grobler (2020) provide empirical findings for the positive impact of ICT on economic growth as well as the impact of ICT on economic development. Albiman and Sulong (2016) find that an Internet penetration rate of above five per cent triggers economic growth and adds that human capital, domestic investments and the quality of institutions are the main transmission channels for ICT usage.

[7] Meaning that a developing country would take significant leaps in their economic development.

3.2 ICT diffusion in relation to education

Previous literature examines the relationship between ICT diffusion and educational outcomes, albeit with varying conclusions.

Bagchi (2005) argues that primary evidence of a positive correlation between the use of ICT and educational attainment exists, controlling for income levels. Furthermore, the study points out a correlation between educational attainment and computer investments, strengthening the argument that higher levels of educational attainment increase ICT diffusion. In a study conducted by Kho, Lakdawala, and Nakasone (2018), the impact of Internet access on the performance of school children is examined within Peruvian schools. Kho, Lakdawala and Nakasone conclude that Internet access in schools positively impacts standardised test results, with an increasing impact over time. A recent study by Bianchi, Song, and Lu (2020) examined a Chinese intervention where teachers were connected with more than 100 million rural children through the use of ICT. Bianchi, Song and Lu concluded that the intervention positively impacted educational performance, use of computers and labour with effects lasting for ten years after the intervention. An empirical study, using panel data for countries within SSA, concludes that enhanced Internet penetration benefits countries with below-median levels of education quality while not being relevant for countries with better than median education quality (Asongu and Odhiambo 2019).

In contrast to these studies, Baliamoune-Lutz (2003) concluded that no relationship between ICT diffusion and education exists in developing countries when regressing ICT indicators on an education index created by the UNDP. In support of this argument, Faber, Sanchis-Guarner, and Weinhardt (2015) studied the relation between ICT availability within children's homes and school test scores and concluded that changes in available broadband connections have zero effect on test scores. Similarly, Fairlie and Robinson (2013) conducted an experimental field study where computers were randomly assigned to children's homes, concluding zero impact on educational performance. A similar study from Peru concluded that assigning children laptops had largely significant impacts on digital skills development, however, no impact on academic performance (Malamud et al. 2018). Bessone, Dahis, and Ho (2020) study the impact of mobile Internet availability on educational outcomes and similarly conclude that the availability of mobile Internet itself does not improve educational performance. Moreover, Porter et al. (2016) discuss the positive and negative aspects of ICT usage in relation to the school environment, arguing for responsible use of ICT to reduce classroom disruptions.

3.3 SCM studies on educational outcomes

The SCM is a relatively new methodology in comparative case studies, but the SCM has nonetheless been used in three studies related to educational enrollment, graduation rates and educational performance, all described below. No previous studies have applied the SCM to empirically study the relationship between ICT and educational attainment[8], but still provide value in regards to the use of the methodology with educational outcomes.

Anghel et al. (2015) studied the impact of standardised school tests on educational performance and leveraged the fact that not all Spanish regions implemented standardised tests in the educational setting. The study makes use of the SCM to strengthen the parallel trends assumption required by the conventional Difference-in-Difference (DiD) methodology but faces limitations due to few pre-intervention observations (Anghel et al. 2015). Cerqua and Di Pietro (2017) studied the effect of a natural disaster on educational enrollment at the local university using the SCM and found a difference in the student population composition due to post-disaster measures. Lastly, Bifulco, Rubenstein, and Sohn (2017) studied the intervention of tuition scholarships and the effect on educational enrollment- and graduation rates in the U.S. using the SCM. The study found that enrollment rates increased, but the weak validity of the results is emphasised due to noisy pre-treatment outcomes when using the SCM (Bifulco, Rubenstein, and Sohn 2017). Despite the varying nature of interventions in SCM studies on educational outcomes, methodology applications regarding similar main outcome variables provide value to this study.

4 Research design and empirical methodology

As presented in section 3.2, several scholars have conducted research on the relationship between ICT and education. Nonetheless, there are several gaps for research to fill. The impact of Internet penetration on educational outcomes is one of those gaps never previously studied using a synthetic control group in a comparative case study. The following section aims to delimit the research gap more clearly, define the research question and present our hypotheses. Furthermore, the chosen empirical methodology will be defined and motivated, and limitations will be outlaid.

[8] To the best of the authors' knowledge at the time of writing this paper.

4.1 Research contribution, research question and hypotheses

This study contributes to the literature through the application of a recently developed empirical method to study the effect of ICT diffusion on children's primary schooling in SSA. Since no studies have utilised the SCM to study the relationship between Internet penetration and educational attainment, our findings add to the existing research on technology diffusion and its impact on children's education. Furthermore, our study specifically advances knowledge on The Gambia's primary school completion rate in relation to the relatively high Internet penetration within the country. In addition, this paper draws upon existing research on the critical threshold of Internet penetration[9] in combination with the SCM to practically apply the SCM in a case where intervention occurs gradually, such as the case with technology adoption.

Moreover, this research contributes to the broader topic regarding the value of investments into ICT development as well as studies on the use of ICT in the educational setting. These contributions might provide value for policymakers in relation to ICT deployment and developments of the education system. Many countries in SSA, as well as other developing countries, face problems related to the digital divide as well as inequalities of educational attainment and illiteracy. Potentially leveraging positive externalities of broad Internet penetration could thus add value to both debates and be a potential way forward for developing countries in general. The main research question that our study aims to answer is:

What are the effects of broad Internet penetration on the national-level primary school completion rate in The Gambia?

With previous literature in mind, our hypothesis is that widespread Internet penetration will contribute positively toward educational attainment. Time lag effects of ICT diffusion need to be considered, however, and results might thus differ depending on the time frame studied.

4.2 The Synthetic Control Method

To derive the causal effect of an intervention, the DiD methodology is applied widely in comparative case studies. The DiD method's ability to separate the causal effect of intervention from other covariates, reasonably affecting the main outcome variable, makes it suitable for case studies as long as the methodology assumptions are fulfilled. The main assumption for the DiD

[9] The critical mass of Internet users needed to ensure exponential network growth as well as continued positive network externalities, further explained in section 4.4.

framework is the parallel trend assumption, meaning that in the absence of the intervention - the treatment group and the control group should have a constant difference over time. The Synthetic Control Method was developed as an extension to the DiD methodology, subsequently relaxing the assumption of parallel trends since the control group is artificially constructed. By creating a control group using synthetic weights of units from a donor pool of similar units to the treatment unit, and a set of key predictors, the parallel trend between the treatment- and the control group can be achieved (Abadie, Diamond, and Hainmueller 2010; Abadie and Gardeazabal 2003).

Furthermore, the DiD methodology limits the user to exclude certain covariates that could potentially be affected by the intervention itself, therefore being considered bad control variables. The SCM instead includes covariates when predicting the pre-intervention control group and uses a data-driven methodology of choosing which units out of the donor pool to include in the control group which reduces user selection subjectivity (Abadie and Gardeazabal 2003; Abadie, Diamond, and Hainmueller 2010).

In our paper, the causal effect of Internet penetration on primary school completion rate can be studied by utilising the SCM. Since data availability restricts us to national-level panel data provided by the World Bank, the SSA, with varying degrees of national Internet penetration, provides a favourable sample for a comparative case study. One of the countries with above-average Internet penetration in SSA is The Gambia, with higher levels of ICT diffusion than many other countries within the macro-region.

4.2.1 Definition

Drawing on Leroutier (2022), the Synthetic Control Method (SCM) will be defined in the following section. By using traditionally used notation for policy evaluation, the estimation of β_{GAt} when $t \geq 2014$, is defined as

$$\beta_{GAt} = Y^1_{GAt} - Y^0_{GAt} \tag{1}$$

where Y^1_{GAt} designates The Gambia's primary school completion rate with above-threshold Internet penetration[10] at each period. Y^0_{GAt} designates The Gambia's primary school completion rate without above-threshold Internet penetration at each period. β_{GAt} measures the difference between the two aforementioned. Estimating β_{GAt} proves difficult since Y^1_{GAt} can be observed when $t \geq 2014$ but Y^0_{GAt} can not be observed.

Following Abadie, Diamond, and Hainmueller (2010), let us assume that the outcome in the absence of above-threshold Internet penetration Y^0_{GAt}, for each country c and period t, can be modelled using the following linear factor model:

$$Y^0_{ct} = \delta_t + Z_c \alpha_t + f'_t \lambda_c + \epsilon_{ct} \tag{2}$$

where δ_t designates a time fixed effect, Z_c designates a vector of observed exogenous country characteristics, α_t is a vector of the unrecognised parameters, f_t designates a vector of unrecognised common factors (and f'_t denotes its' transpose), λ_c designates a vector of unrecognised country-level effects or factor loadings, and ϵ_{ct} designates the error term with a mean of zero (eg. capturing transitory shocks at the national-level).

Since the linear factor model above allows for time effects and individual effects to interact, it is more flexible than the typical DiD framework. Abadie, Diamond, and Hainmueller (2010) prove that it is possible to use β_{GAt} as an estimator for the function of outcomes observed in the post-treatment in other countries, with these specifications:

$$\widehat{\beta}_{GAt} = Y_{GAt} - \sum_{j=1}^{J} \omega^*_j Y_{jt} \tag{3}$$

[10] Above-threshold Internet penetration is defined in section 4.4.

where $\sum_{j=1}^{J} \omega_j^* Y_{jt}$ denotes the weighted combination of the outcome for J countries not exposed to above-threshold Internet penetration, and where the vector $W^* = (\omega_1^* ...... \omega_J^*)'$ should satisfy the four following conditions:

$$\begin{cases} \omega_j^* \geq 0 \forall j = 1..J \\ \sum_{j=1}^{J} \omega_j^* = 1 \\ \overline{Y}_{GA}^K = \sum_{j=1}^{J} \omega_j^* \overline{Y}_j^K \\ Z_{GA} = \sum_{j=1}^{J} Z_j \end{cases} \tag{4}$$

where $\overline{Y}_{GA}^K$ denotes a linear combination of pre-intervention primary school completion rate in The Gambia, and $\overline{Y}_j^K$ a linear combination of pre-intervention primary school completion rate for country j (e.g. the simple mean of pre-intervention outcomes $\overline{Y}_j^K = 1/T_0 \sum_{t=1}^{T_0} Y_j$). Abadie, Diamond, and Hainmueller (2010) further emphasise the importance of a high number of pre-treatment periods, since the estimator gets closer to the true parameter β_{GAt} when the number of pre-treatment periods is high compared to the scale of transitory shocks affecting countries.

Estimating the appropriate vector W in practice is done in STATA by using the command named **synth2** developed by Abadie, Diamond and Hainmueller (2010). The algorithm minimises the distance between a vector of the pre-intervention characteristics in the treated unit, X_{GA}(with dimensions $K\ x\ 1$) and a weighted matrix of pre-intervention characteristics in the non-treated units, $X_0 W$ (with dimensions $K\ x\ K$). Pre-intervention characteristics are of two types; firstly, the linear combinations of pre-intervention outcomes $\overline{Y}_j^K$; and secondly, the country characteristics Z_j not affected by the intervention. To obtain the vector W, the algorithm starts with a positive and semi-defined matrix V that defines a dot-product. The distance between X_{GA} and X_0W is then denoted:

$$X_{GA} - X_0W \; = \; \sqrt{(X_{GA} - X_0W)'V(X_{GA} - X_0W)} \tag{5}$$

where the goal is to estimate the $W^*(V)$ that minimises the distance. Such minimisation results can be shown to be equivalent to a diagonal matrix assigning weights to the linear combination of characteristics in X_{GA} and X_0W. Following Abadie and Gardeazabal (2003), we estimate V by minimising the root mean squared prediction error (RMSPE)[11] of the outcome variable in the pre-treatment periods. Formally, let Y_{GA} be the $(23\ x\ 1)$ vector of pre-2014 primary school completion rate from 1990 to 2013 for The Gambia and Y_j be the $(23\ x\ J)$ matrix of pre-2014 primary school completion rate for the J other SSA countries in the donor pool. V^* is chosen such that:

$$V^* = \; argmin(Y_{GA} - Y_jW^*(V))'(Y_{GA} - Y_jW^*(V)) \tag{6}$$

where V is the set of all non-negative diagonal $(K\ x\ K)$ matrices.

To estimate a satisfactory synthetic control, two criteria need to be fulfilled. Closeness of pre-intervention characteristics, both for the treated and the synthetic unit, is needed. The characteristics' closeness depends on the ability of the characteristics to predict the outcome, which can be assessed by comparing pre-intervention characteristics for the treated unit and the synthetic unit. Secondly, closeness of pre-intervention outcomes, both for the treated and the synthetic unit, is needed. The outcome closeness can be assessed graphically or by computing the RMSPE.

4.2.2 Inference

Assessing the uncertainty in SCM estimates using asymptotic approximation proves difficult since estimates draw on the comparison between a single treated unit and the synthetic

[11] The RMSPE gives the average of the root squared difference between the treated unit's- and the synthetic control's pre-intervention outcomes.

counterpart (Bifulco, Rubenstein, and Sohn 2017). Instead of using asymptotic approximations, Abadie, Diamond, and Hainmueller (2010) draw on permutation tests to statistically assess if the outcome effect for the treated unit is sufficiently unlikely to happen at random. By constructing a Leave-One-Out (LOO) robustness test, where a synthetic control is created for each unit of the sample (with the remaining units as the donor pool), it is possible to assess the likelihood of an outcome effect happening by chance. Since units in the LOO robustness test are untreated to the intervention, the outcome effect can be presumed to be zero. Comparing the outcome effects in the robustness test with the actual outcome for the treated unit, the likelihood of an outcome effect happening by chance becomes apparent (Bifulco, Rubenstein, and Sohn 2017; Abadie, Diamond, and Hainmueller 2010).

4.3 Selection of predictors

In order to create the counterfactual unit, different weights are assigned to different units within the donor pool using predictor variables. Selecting these predictors is an essential step of the methodology in order for the synthetic unit to minimise the RMSPE. The selection of predictors naturally includes the main outcome variable for the pre-intervention period, which helps control for unobserved factors (Abadie, Diamond, and Hainmueller 2010). The inclusion of the main outcome variable as a predictor does not guarantee, however, that the synthetic control will result in a close match to the treatment country. Since the SCM is relatively new, existing literature does not provide thorough guidance on the selection of appropriate predictors (Bifulco, Rubenstein, and Sohn 2017). "Once it has been established that the unit representing the case of interest and the synthetic control unit have similar behavior over extended periods of time prior to the intervention, a discrepancy in the outcome variable following the intervention is interpreted as produced by the intervention itself" (Abadie, Diamond, and Hainmueller 2010).

Since previous literature lacks thorough guidance on the choice of predictor variables, our study will follow other studies with similar main outcome variables, related to educational outcomes, using the SCM (cf. Anghel et al. 2015; Cerqua and Di Pietro 2017; Bifulco, Rubenstein, and Sohn 2017).

4.4 The critical threshold of Internet penetration

Since our study is based on the intervention of Internet usage itself, which of course is a gradual process of adoption within user groups and countries globally, the delimitation of countries available for a donor pool needs further definition. Since the SCM requires that donor pool units are untreated for the intervention being studied, a binary definition of considerable Internet penetration is needed. As aforementioned, the digital divide is particularly present in SSA, with a small fraction of the population using the Internet in many countries. The definition of what could be considered a threshold for Internet penetration has been discussed by several scholars previously. Since the Internet acts as the prime example of a network effect[12], previous literature tries to define the critical mass[13] of Internet users. Reaching a critical mass of Internet users is considered to be the tipping point of the gradual Internet penetration within a group of users, which subsequently ensures both widespread Internet penetration within a period of time as well as continued network effects. The diffusion of technology[14] gives rise to positive network externalities for the users, which in turn further increases the technology diffusion, and this positive loop leads to exponential diffusion of the network. The exponential growth of the network, however, can only be achieved if the society manages to reach the critical mass threshold required for self-perpetuating network growth (Lechman 2015).

Combining the SCM with a theoretical framework for critical masses for positive network externalities allows for an application of the SCM in a case study where the intervention itself happens gradually. While several studies have been conducted to define the critical threshold of Internet penetration (cf. Wamboye, Tochkov, and Sergi 2015; Albiman and Sulong 2016; S. Dutta et al. 2009; Gillwald and Mothobi 2019; Asongu and Acha-Anyi 2020; Lechman 2015), there is a lack of general consensus regarding the exact critical threshold level. The critical threshold is within the range of 4,5 - 20 per cent of the population using the Internet, according to all studies. The Global Information Technology Report (GITR)[15] classifies several critical mass thresholds accelerating network connectivity. The 'familiarization'-phase of Internet penetration, according to the GITR, is defined as countries with at least 15 per cent of the population using the Internet, and this phase is considered to create Internet momentum (S. Dutta et al. 2009). Asoungu and Acha-Anyi (2020) studied the impact of ICT diffusion on economic productivity in SSA and concluded that the level of 15 per cent Internet penetration constitutes a critical mass

[12] According to the Oxford Dictionary, the effect that happens when a product or service gains value as more people use it.

[13] According to the Oxford Dictionary, the minimum amount of resources, number of customers, etc. needed to start or support a project or an activity, or the minimum size that a project or activity needs to be in order to be successful.

[14] "Technology diffusion, which is defined as a dynamic and time-attributed process involving the transfer of information, knowledge and innovations, and standing for a continuous and gradual spread of new ideas throughout large-scale and heterogeneous societies" (Lechman 2015).

[15] A report produced in cooperation by INSEAD, World Economic Forum, and the World Bank, the Global Information Technology Report ranks 144 economies in terms of networked readiness (ICT) and its effects on, economic growth and productivity.

threshold for increased economic productivity. Based on the comprehensive and well-reputed nature of the GITR, our study will follow the 15 per cent 'familiarization'-phase critical threshold for Internet penetration outlined in GITR 2008-2009, also supported by the 15 per cent threshold presented by Asongu and Acha-Anyi (2020) studying SSA in particular.

4.5 Assumptions and limitations

Using the SCM, there are limitations to consider, which are especially important when deciding upon predictor variables and donor pool countries. The first essential limitation is to ensure that no members of the donor pool have been exposed to the intervention. In the case of SGA, no donor pool countries should thus be affected by Internet penetration above the critical mass threshold since that could lead to skewness in the outcome variable. Furthermore, the SCM risks interpolation bias if members of the donor pool differ in characteristics from the treatment unit (Abadie 2021). Only countries with similar characteristics as The Gambia should thus be part of the donor pool. The delimitation of donor pool countries to SSA, with similar factors of economic development, geopolitics, history, culture and sociology (Anděl, Bičík, and Bláha 2018), thus becomes important to reduce the risk of interpolation bias.

Idiosyncratic shocks to the major outcome variables affecting countries in the donor pool also have to be considered to minimise the limitations of the SCM (Abadie 2021). The selection process for donor pool countries thus includes qualitative research of educational policy changes in SSA as well as other potential idiosyncratic shocks, and subsequently excludes countries exhibiting such idiosyncratic shocks from the donor pool.

In addition, the SCM requires a long pre-treatment period since the validity of the method is based on a pre-treatment path and pre-treatment level that is a close match between the treatment unit and the synthetic control unit (Abadie, Diamond, and Hainmueller 2015). Since the pre-treatment period for the creation of SGA stretches over a period of 23 years, the requirement of a long pre-treatment period should be fulfilled. Furthermore, if an intervention causes changes over time, a longer post-treatment period is required for the appropriate use of the SCM (Abadie, Diamond, and Hainmueller 2015). The post-treatment period for SGA is limited to a period of four years due to missing data after 2017 as well as several donor pool countries reaching the critical threshold of Internet penetration, subsequently excluding them from the donor pool. The post-treatment period of only four years should be kept in mind when analysing the results since a longer post-treatment period would be beneficial to draw conclusions regarding the long-term effects of the intervention.

Moreover, limitations of the 15 per cent critical mass threshold for Internet penetration have to be considered in regard to the empirical methodology. Acknowledging that other scholars conclude that other levels of Internet penetration constitute the critical threshold in SSA (cf. Wamboye, Tochkov, and Sergi 2015; Albiman and Sulong 2016), the exact threshold level of 15 per cent is not widely accepted. Wamboye, Tochkov, and Sergi (2015) as well as Albiman and Sulong (2016) find the critical threshold to be approximately five per cent of a population using the Internet in SSA. These thresholds are in the lower end of the 4.5 - 20 per cent interval, however, and Gillwald and Mothobi (2019) find the critical threshold to be 20 per cent. Since our study uses the critical threshold of 15 per cent, potential limitations following extreme threshold values at the different ends of the interval might be avoided. Furthermore, while different threshold levels could potentially cause the SCM results to be more or less robust, The Gambia has nonetheless been exposed to higher levels of Internet penetration than the donor pool. Choosing another threshold level would not cause that relationship to change but would justify a different selection of intervention year, pre-intervention period and post-intervention period.

5 Data

In order to use the Synthetic Control Method to study the causal effect of above-threshold Internet penetration on primary school completion rate, three different types of panel data are collected for all countries where data is available in SSA. Firstly, data on Internet usage is collected for all SSA countries to determine the donor pool. Secondly, data on educational outcomes in countries within the donor pool is collected. Lastly, data on the selected predictor variables are collected for the treatment country as well as the countries within the donor pool. All variables used are in relation to the population size, either by default or adjusted accordingly. The summarised data is available in Appendix A. Collected data is moreover divided into pre-intervention data and post-intervention data. The selection of donor pool countries is furthermore defined, and limitations are presented.

5.1 Selection of intervention year

Based on the selected 15 per cent threshold level of Internet penetration, 2014 is the year when The Gambia surpassed the critical threshold level, which will be considered the intervention year.

The pre-intervention period thus stretches from 1990 until 2013, and the post-intervention period stretches from 2014 up until 2017[16].

5.2 Donor pool countries

To make sure that SGA does not get affected by above-threshold Internet penetration in donor pool countries, an exclusion of countries in SSA that are not suitable for the donor pool has been made. Firstly, in accordance with section 2.1, the study only includes countries in SSA since countries within the same macro-region are considered comparable. The Republic of South Sudan (South Sudan) and the Republic of Sudan (Sudan) have been excluded due to the recent independence of South Sudan on July 9 2011 (United States Institute of Peace n.d.). Due to the difficulties of separating data for the countries respectively, these two countries are excluded from the donor pool.

After these exclusions were made, all countries with an Internet penetration exceeding the 15 per cent threshold are considered to be treated by the intervention. All countries at the end of the post-intervention period, in the year 2017, exceeding an Internet penetration of 15 per cent are thus excluded from the donor pool. All remaining countries, not exceeding the 15 per cent threshold between 2014 and 2017 can thus be included in the donor pool for SGA. Due to a general lack of data availability, some additional countries had to be excluded from the donor pool. The following countries were excluded due to a lack of data; the Central African Republic, the Republic of the Congo (Congo), the Union of Comoros (Comoros), the Republic of Malawi (Malawi), the Republic of Sierra Leone (Sierra Leone), the Federal Republic of Somalia (Somalia), the Republic of Uganda (Uganda), and the Republic of Zambia (Zambia).

In accordance with Section 4.5, the available remaining donor pool countries were researched in regard to potential idiosyncratic shocks and major policy changes. Out of the remaining available donor pool countries, the Republic of Burundi (Burundi) showed signs of an idiosyncratic shock in 1993 whereafter the primary school completion rate dropped heavily (see Figure 3). In October 1993, the first democratically elected president was assassinated in a coup attempt which subsequently led to the start of the civil war in Burundi (Verwimp and Van Bavel 2014). Verwimp and Van Bavel (2014) further highlight that the primary school completion rate dropped by almost four percentage units, significant at the one per cent level, for children

[16] Data availability is lacking for several countries after 2017 in combination with several donor pool countries reaching the critical threshold of 15 per cent Internet penetration thereafter.

exposed to the violent conflict in Burundi. Following Abadie (2021), donor pool units with idiosyncratic shocks affecting the main outcome variable should be excluded to minimise the limitations of the SCM. Burundi is thus excluded from the donor pool in our study.

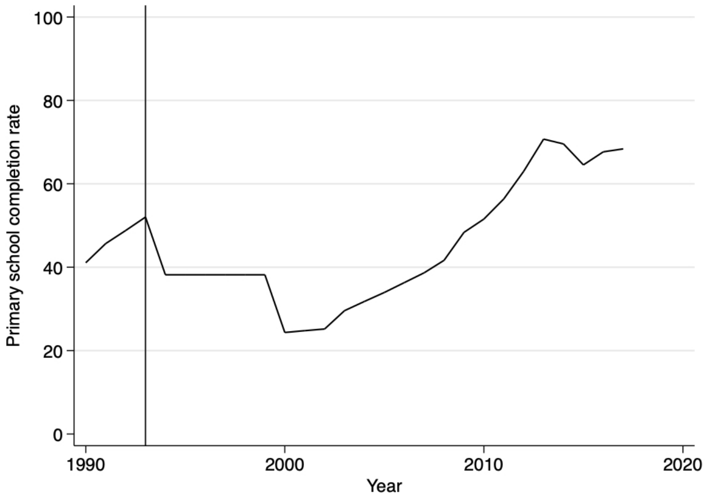

Figure 3: Primary school completion rate in Burundi

Source: Data for 1990-2017 provided by the World Bank

The remaining countries that are included in the final donor pool for Synthetic The Gambia are; Burkina Faso, the Republic of Chad (Chad), the State of Eritrea (Eritrea), the Republic of Madagascar (Madagascar), the Republic of Mozambique (Mozambique), the Republic of the Niger (Niger), and the Togolese Republic (Togo).

5.3 Variables

The dataset used in the study is sourced from the World Bank through the STATA-connected database called “wbopendata”[17]. The database is built on the main World Bank collections of different development indicators. The World Bank, in turn, sources data from different official international institutions. Firstly, all indicators within the topic of infrastructure were gathered to gain access to our intervention variable for all countries in SSA. Subsequently, all indicators within the topic of education were collected to gain access to our main outcome variable and different predictor variables used in the study.

[17] Available for download through https://datahelpdesk.worldbank.org/knowledgebase/articles/889464-wbopendata-stata-module-to-access-world-bank-data

5.3.1 Main outcome variable

Primary school completion rate: The main outcome variable used to measure educational outcome is the primary school completion rate as a percentage of the relevant age group. According to the World Bank (n.d.b), the variable is frequently being used as a core indicator for the performance of an education system. The variable is measured by the number of students enrolled minus repeaters in the last year of primary school, divided by the population of students in the age group which should attend the last year of primary school.

5.3.2 Intervention variable

Individuals using the Internet: The intervention variable, which delimits which countries can be included within the donor pool for SGA, measures the percentage of individuals using the Internet in relation to the population size. The data is based on how many individuals used the Internet by computer, mobile, digital assistant, game machines, television, et. cetera. during the last three months.

5.3.3 Predictor variables

To construct SGA, six different predictor variables are selected using the methodology explained in section 4.3.

Primary school completion rate: Firstly, the lagged main outcome variable, primary school completion rate, will be used as an important predictor variable. According to McClelland and Gault (2017), using the lagged main outcome variable avoids the potential omitted predictors' effect since the lagged main outcome variable takes potentially omitted predictor variables into account. Including the lagged main outcome variable as a predictor for only some of the pre-treatment years is conventional practice since including all pretreatment years would result in an elimination of the effect of all other predictor variables (McClelland and Gault 2017).

Population aged 0-14: The percentage of individuals aged between 0 and 14 years old in relation to the total population will be used to account for differences in the age group, mostly relevant to the primary school completion rate. Cerqua and Di Pietro (2017) have similarly used the proportion of individuals aged below 20 in their SCM study on educational outcomes.

Pupil-teacher ratio primary school: The pupil-teacher ratio for primary education is used as a predictor since a higher ratio might lead to better education system performance, and the same predictor variable is used in the SCM study on educational outcomes by Anghel et al. (2015).

Probability of children dying: The probability of children between the ages of 5-9 and 10-14 dying will be used as predictors since it accounts for the health status of the relevant age group as well as an indicator of the country's socioeconomic status. Pieters et al. (2014) conducted an SCM study on child mortality where primary school completion rate was used as a predictor variable, strengthening the argument for an important connection between the two variables.

GDP per capita, constant 2010 (US dollars): Lastly, GDP real per capita will be added as a predictor of a country's overall economic development during the pre-intervention period, and GDP per capita as a predictor variable has similarly been used in the SCM study by Cerqua and Di Pietro (2017) on educational outcomes.

5.4 Data limitations

The main data limitation in this study has been missing observations for a number of variables, countries and periods. To resolve the issue of missing observations, mean imputation has been conducted to replace missing observations with mean values of the closest observations. Countries considered to have too many missing data observations were excluded from the selection in section 5.2. The percentage of variable samples where missing values have been replaced using mean imputation is presented in Table 1.

Table 1: Mean imputation

Variable	% of Sample
Primary school completion rate	13,1%
Pupil-teacher ratio	4,8%
Probability of children dying, aged 5-9	3,2%
Probability of children dying, aged 10-14	3,2%
Population aged 0-14	2,0%

Source: Data provided by the World Bank for the period of 1990-2017, with mean imputation conducted by the authors

Furthermore, Eritrea missed observations for *GDP per capita, constant 2010 (US dollars)* for the period of 2012-2017 as well as the period 1990-1991. These missing values have been replaced by

GDP data provided by the International Monetary Fund (IMF) database[18]. Real GDP growth (%) for Eritrea has been used to manually calculate missing values, using the latest available observation from the World Bank dataset.

6 Results

The empirical results from the study are presented in the section below. Firstly, the result of the data-driven construction of Synthetic The Gambia is presented with corresponding donor pool country weights. Secondly, the comparison between The Gambia and Synthetic The Gambia in regards to primary school completion rate is presented. Lastly, the results from several placebo tests and one robustness test on the primary school completion rate are presented.

6.1 Constructing Synthetic The Gambia

Following the empirical methodology explained in section 4.2, the weights of each country in the donor pool are calculated. The construction of SGA requires choosing a combination of predictor variables resulting in the lowest possible RMSPE. Minimising the RMSPE will ensure that both SCM closeness criteria are fulfilled, namely the closeness of characteristics pre-intervention and the closeness of outcomes pre-intervention.

6.1.1 Selecting predictor variables

According to section 4.3, the predictor variables are chosen through a literature-driven approach. Subsequently, the primary school completion rate is lagged, and the main outcome variable is included as a predictor variable. All possible combinations of the other predictor variables are then tested using a data-driven approach with the goal of minimising the RMSPE. The comparison of different predictor combinations is presented in Table 2, where the predictor combination in row 20 results in the lowest RMSPE. The best combination includes seven predictors; five predictors of the lagged primary school completion rate, the probability of children (5-9 years old) dying and the probability of children (10-14 years old) dying. For this combination of predictors, the RMSPE is 7.68753, which is the mean difference between GA and SGA in the pre-intervention period.

[18] Available at https://www.imf.org/en/Data.

Table 2: Test of predictor combinations

	Primary school completion rate	Primary school completion rate	Primary school completion rate	Primary school completion rate	Primary school completion rate	Population aged 0-14	Pupil-teacher ratio primary school	Probability of children dying 5-9	Probability of children dying 10-14	GDP per capita, contant 2010 (US dollar)	
Test #	**1992**	**1997**	**2002**	**2007**	**2012**	**1990-2013**	**1990-2013**	**1990-2013**	**1990-2013**	**1990-2013**	**RMSPE**
1	0.0835	0.0000	0.1458	0.0000	0.2404	0.0705	0.0875	0.0922	0.0946	0.1855	10.53409
2	0.1010	0.0000	0.3106	0.3848	0.0950	-	0.0055	0.0517	0.0432	0.0081	7.76621
3	0.0409	0.2430	0.1213	0.4106	0.0358	-	-	0.0685	0.0776	0.0022	7.70605
4	0.0903	0.1367	0.2885	0.3783	0.0742	-	-	-	0.0157	0.0163	7.77055
5	0.1025	0.1384	0.3355	0.3253	0.0837	-	-	-	-	0.0146	7.76585
6	0.0948	0.0000	0.2840	0.4039	0.0811	0.0046	-	0.0600	0.0399	0.0316	7.96883
7	0.0831	0.1245	0.2769	0.3957	0.0709	0.0027	-	-	0.0235	0.0227	7.91556
8	0.1034	0.1790	0.3353	0.2455	0.0870	0.0364	-	-	-	0.0134	7.85867
9	0.1151	0.2163	0.2795	0.2723	0.0814	0.0355	-	-	-	-	7.79472
10	0.0884	0.1403	0.2598	0.0000	0.1995	0.0352	0.0629	-	0.1251	0.0889	9.54377
11	0.1020	0.1425	0.3451	0.2911	0.0967	0.0012	0.0087	-	-	0.0128	7.77364
12	0.0333	0.2949	0.1960	0.1349	0.0717	0.1985	0.0708	-	-	-	8.80294
13	0.1091	0.1751	0.2996	0.3112	0.0927	-	0.0124	-	-	-	7.73326
14	0.1017	0.0000	0.2815	0.2683	0.1181	0.0224	0.0245	0.1141	-	0.0693	8.44411
15	0.1028	0.1425	0.3314	0.2977	0.0904	0.0028	0.0115	0.0210	-	-	7.74846
16	0.0990	0.1314	0.3476	0.2927	0.0893	-	0.0182	0.0219	-	-	7.74889
17	0.1084	0.2066	0.2836	0.2940	0.0903	-	-	0.0172	-	-	7.74285
18	0.1012	0.0000	0.3074	0.3722	0.0994	0.0030	0.0137	0.0540	0.0491	-	7.77095
19	0.0983	0.0000	0.3167	0.3676	0.0976	-	0.0194	0.0573	0.0430	-	7.77400
20	**0.0304**	**0.2985**	**0.0716**	**0.3710**	**0.0256**	**-**	**-**	**0.1113**	**0.0916**	**-**	**7.68753**
21	0.0901	0.2350	0.2327	0.3611	0.0688	-	-	-	0.0123	-	7.74141
22	0.0900	0.1891	0.3172	0.2272	0.0727	0.0070	-	0.0537	-	0.0433	8.12676
23	0.0881	0.1518	0.3151	0.3293	0.0791	0.0029	0.0133	-	0.0203	-	7.76764
24	0.0340	0.2916	0.0846	0.3763	0.0282	0.0004	-	0.1007	0.0843	-	7.68963
25	0.0933	0.1425	0.3226	0.3475	0.0773	-	0.0042	-	0.0051	0.0077	7.76544
26	0.1019	0.1407	0.3541	0.2936	0.0896	-	0.0057	0.0060	-	0.0084	7.76498
27	0.1071	0.1879	0.2901	0.2679	0.0840	0.0441	-	0.0190	-	-	7.81794
28	0.0968	0.1544	0.2608	0.3498	0.0719	0.0490	-	-	0.0173	-	7.83863
29	0.0816	0.1663	0.3297	0.3017	0.0789	-	0.0210	-	0.0208	-	7.76378
30	0.1021	0.1358	0.3369	0.3018	0.0996	-	0.0118	-	-	0.0121	7.76986
31	0.1048	0.1526	0.3381	0.2644	0.0841		-	0.0302	-	0.0258	7.84805

Source: Data for 1990-2017 provided by the World Bank and missing values for Eritrea are replaced with data provided by the International Monetary Fund

6.1.2 Comparison of outcomes for SGA and GA

If the primary school completion rate for SGA follows the curve for GA closely pre-intervention, then the two closeness criteria of the SCM are considered fulfilled. The difference between SGA and GA is, on average, 7.69 percentage points in the pre-intervention period. As visible in Figure 4, there is a visible trend that SGA's outcome follows that of GA. The plotted lines tend to differ at times, and even though the difference is larger during some pre-intervention years, we conclude that the plotted lines follow each other more closely during the pre-intervention period than during the post-intervention period. The difference between the primary school completion rates of SGA and GA is larger in the post-intervention period than at any time during the pre-intervention period.

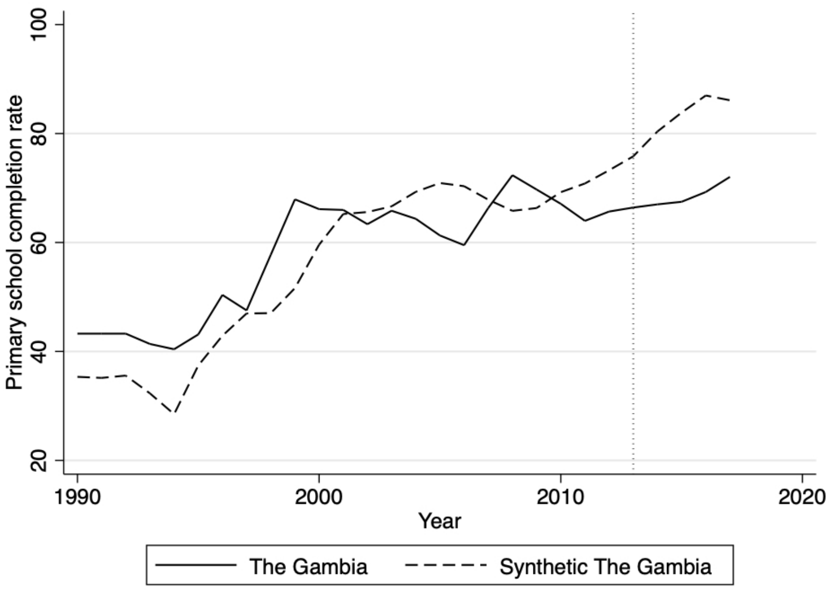

Figure 4: Comparison of GA and SGA

Source: Data for 1990-2017 provided by the World Bank

Additionally, in order to validify the chosen selection of predictor variables for SGA, the predictor values for SGA are compared with the average control values, see results in Table 3. The average control (AC) denotes the mean for the respective predictor variable, where all countries within the donor pool are included with equal weights. The differences, in percentage values, between the predictor values of SGA and GA are presented in column five (SCM Bias). The differences in percentage values between the predictor values of GA and the AC values are presented in column seven (AC Bias). The SCM Bias percentage value is the smallest for the primary school completion rate (1997, 2007) and the probability of children (aged 10-14) dying, meaning that those predictor variables produce the best estimates for the actual outcome of GA. The SCM Bias percentage values for primary school completion rate (1992, 2002) and the probability of children (aged 5-9) dying greatly outperform the AC Bias percentage values, meaning that the constructed SGA has a greater fit than the average mean within the donor pool. Using the SCM for our study thus produces a better estimation than an average control of the donor pool.

The assigned predictor weights are furthermore presented in Table 3, with the predictor variables in column one and their respective weights in column two. The SCM Bias percentage value for primary school completion rate (2012 lag) is nearly as large, in absolute terms, as the AC Bias percentage value resulting in a very low assigned predictor weight. In addition, 79.71 per cent of total predictor weights are assigned to lagged primary school completion rate variables.

According to McClelland and Gault (2017), the lagged main outcome variable is typically the best predictor within an SCM study, and the level of 79.71 per cent of total predictor weights for the lagged main outcome variable is in line with the level in the study by Abadie, Diamond, and Hainmueller (2010).

Table 3: Predictor values pre-treatment

Predictor	Weight	The Gambia	Synthetic The Gambia	SCM Bias	Average Control	AC Bias
Primary school completion rate 1992	0.0304	43.2734	35.5749	-17.79%	26.1705	-39.52%
Primary school completion rate 1997	0.2985	47.5431	46.9784	-1.19%	29.4214	-38.12%
Primary school completion rate 2002	0.0716	63.3650	65.5911	3.51%	36.8528	-41.84%
Primary school completion rate 2007	0.3710	66.3371	67.8578	2.29%	51.6441	-22.15%
Primary school completion rate 2012	0.0256	65.6831	73.2624	11.54%	57.4212	-12.58%
Probability of children dying 5-9	0.1113	14.4250	17.6742	22.52%	22.3238	54.76%
Probability of children dying 10-14	0.0916	8.0750	8.2750	2.48%	11.0238	36.52%

Note: Weight is the optimal covariate weight in the diagonal of the V matrix. The Gambia predictor values denote the simple average for The Gambia. Synthetic The Gambia is the weighted average of control units in the donor pool with optimal weights. SCM Bias is the comparison between The Gambia and Synthetic The Gambia. Average Control is the simple average of control units in the donor pool with equal weights. AC Bias is the comparison between The Gambia and the Average Control.

Source: Data for 1990-2017 provided by the World Bank

Furthermore, the assigned weights of the countries within the donor pool are presented in Table 4. The SCM assigns most of the weight to Togo (78.30%), which can be explained by the fact that Togo displays similar characteristics to those of The Gambia. Togo has a higher primary school completion rate than The Gambia during the period of 2000 to 2013, when most other donor pool countries have lower primary school completion rates, see Figure 9 in Appendix B. Madagascar displays similar trends as The Gambia from the year 2000 onwards with lower outcome values prior to the year 2000. Burkina Faso displays lower outcome values than The Gambia for the entire duration of the pre-intervention period. The combination of these three donor pool countries thus provides a great combination for the predicted path of SGA.

Table 4: Country weights for SGA

Donor pool countries	Weight
Burkina Faso	0.0480
Chad	0.0000
Eritrea	0.0000
Madagascar	0.1690
Mozambique	0.0000
Niger	0.0000
Togo	0.7830

Source: Data for 1990-2017 provided by the World Bank

To conclude, predictor variable weights and donor pool country weights construct the predicted outcome for SGA. The predicted pre-intervention path for SGA is noisier than expected, not closely resembling the actual outcome of GA. The results have fulfilled the criteria of the SCM, however, and the estimation can thus be used to study the effect on primary school completion rate.

6.2 Treatment effect on primary completion

The difference in primary school completion rate between GA and SGA shows the effect of above-threshold Internet penetration on primary school completion rate. Figure 5 displays the difference in outcome for GA and SGA pre-intervention and post-intervention. Prior to the intervention year, the goal is for the treatment effect to be as close to zero as possible. The treatment effect ranges from approximately -10 to +10 per cent during the entire pre-intervention period, with the exception of some minor deviations in 1994 and 2006 and one larger deviation in 1999.

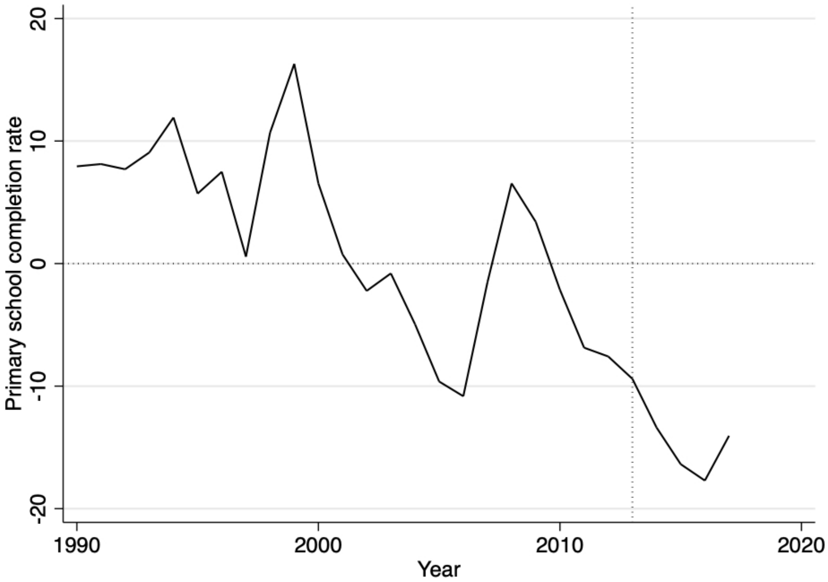

Figure 5: Treatment effect

Source: Data for 1990-2017 provided by the World Bank

Post-intervention, the treatment effect is largely negative for the entire period. The results of the intervention are presented in Table 5, and the average treatment effect post-intervention is -15.38 percentage points with an average negative change of 20.71 per cent. The treatment effect is increasingly negative for all years up until 2017, in which the treatment effect is decreasingly negative. The difference in 2017 might point to differing short-term and long-term effects, but since the post-intervention period only stretches over a period of four years, it is difficult to draw any conclusions based solely on this. The negative treatment effect post-intervention thus means that above-threshold Internet penetration has a negative effect on primary school completion rate, which stands in contrast to our prior hypothesis.

Table 5: Treatment effect on primary completion

Year	Actual outcome	Predicted outcome	Treatment effect	Treatment effect %
2014	66.9994	80.3754	-13.3759	-19.96
2015	67.4725	83.8425	-16.3700	-24.26
2016	69.2726	86.9795	-17.7069	-25.56
2017	72.0460	86.1020	-14.0560	-19.51
Mean	**68.9476**	**84.3248**	**-15.3772**	**-20.71**

Note: The average treatment effect during the post-treatment period is -15.3772

Source: Data for 1990-2017 provided by the World Bank

6.3 Placebo tests and robustness tests

In accordance with section 4.2.2, the results need further validation through the creation of several placebo tests and a robustness test. The tests are conducted to ensure that the treatment effect actually depends on above-threshold Internet penetration, and all tests follow the approach by Abadie, Diamond, and Hainmueller (2015). Firstly, an in-time placebo test is conducted to see if a change of intervention year would affect the results. Secondly, an in-space placebo test is conducted to see if a change of treatment country would affect the results. Lastly, a LOO robustness test to see if the results are affected by one of the donor pool countries included in the construction of SGA.

6.3.1 In-time placebo test

The in-time placebo test is generated by artificially setting the intervention year to an earlier year than the actual intervention year. A manipulated result is then created using the new artificial intervention years. The in-time placebo test is conducted using three different intervention years; 2000, 2005 and 2010. The results are presented in Figure 6 and Table 6.

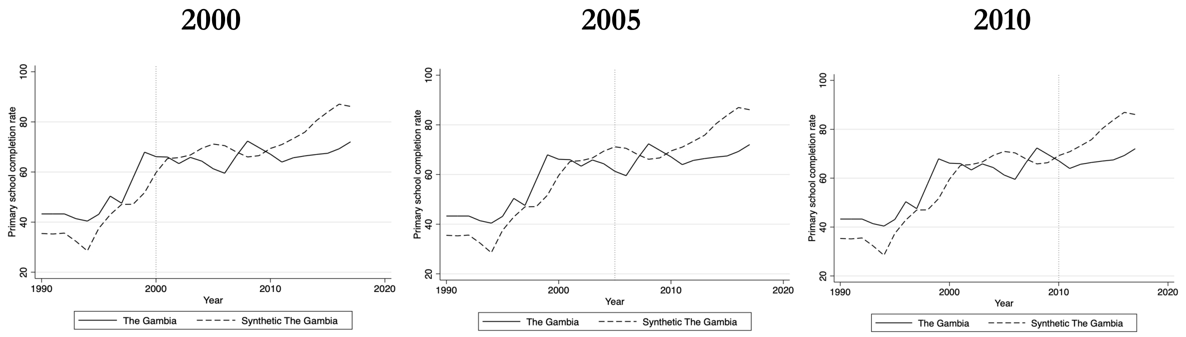

Figure 6: In-time placebo tests

Source: Data for 1990-2017 provided by the World Bank

For all the in-time placebo tests, average treatment effects are negative but not to the same extent as when using the actual intervention year. The divergence of treatment effects might be due to several factors. Firstly, the critical threshold for Internet penetration might need adjustment since ICT diffusion might impact primary school completion rates earlier than expected. Secondly, other factors than Internet penetration could potentially explain the decrease in primary school completion rate in The Gambia, subsequently lowering the validity of our results.

Table 6: In-time placebo tests

Year	Treatment effect
2000	-6.4200
2005	-7.6347
2010	-12.1494

Note: All in-time placebo tests are negative for the post-treatment period.

Source: Data for 1990-2017 provided by the World Bank

6.3.2 In-space placebo test

The in-space placebo test allocates the intervention to another country than the treatment country and repeats the test for all countries within the donor pool. If other countries than the treated country display similar treatment effects to that of GA, other factors affect the primary school completion rate negatively since none of the donor pool countries has been exposed to above-threshold Internet penetration.

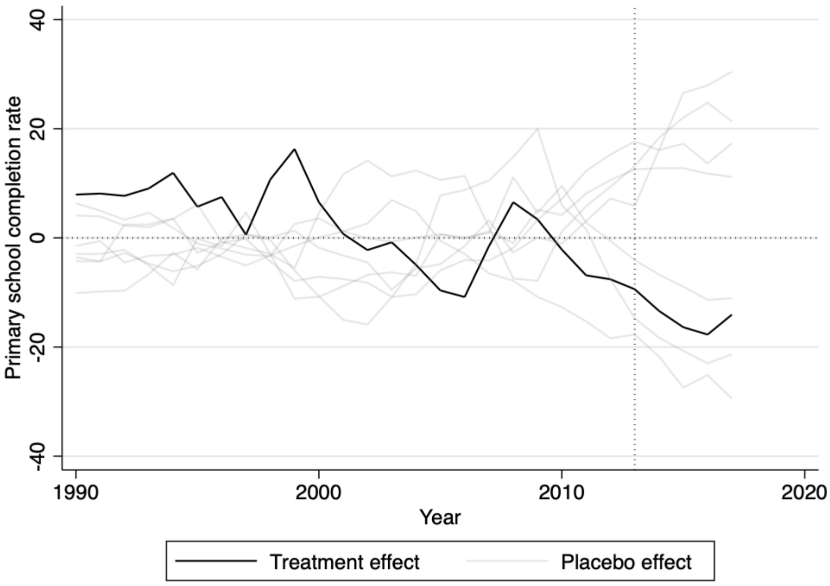

Figure 7: In-space placebo test

Source: Data for 1990-2017 provided by the World Bank

Figure 7 displays the results of the in-space placebo test. The test shows that The Gambia is not the only country with a negative treatment effect on the primary school completion rate. The in-space placebo test was repeated, excluding countries with two times the RMSPE[19], a conventional practice for SCM studies, but the results are similar to the ones presented in Figure 7.

6.3.3 Leave-One-Out robustness test

To further validify the results in section 6.2, a LOO robustness test is conducted. The test evaluates if the main outcome variable is affected by one or a few other donor pool countries. Figure 8 and Table 7 display the test results in the case that one of the positively weighted donor pool countries (Togo, Madagascar, or Burkina Faso) was left out of the sample. If Togo is excluded from the sample, the treatment effect on primary school completion rate would have been positive. With Togo included in the sample, the results display the opposite effect. Togo thus affects the SCM estimation negatively, weakening the validity of our results.

[19] See Figure 10 in Appendix B.

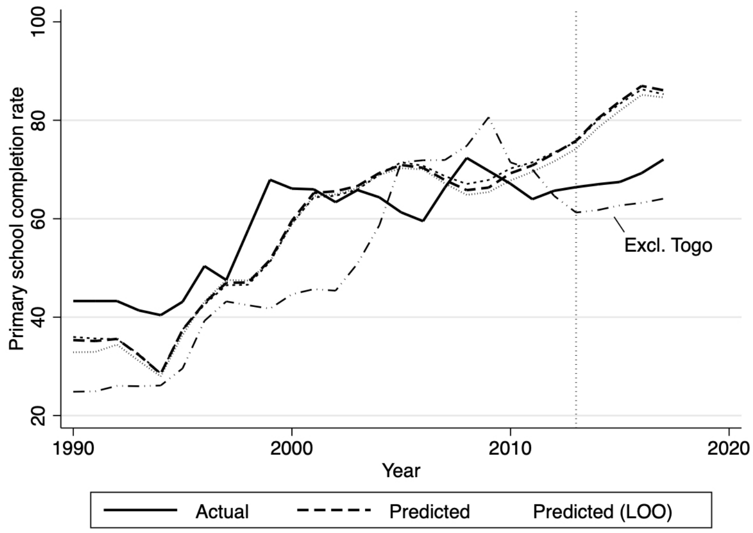

Figure 8: Leave-One-Out robustness test

Source: Data for 1990-2017 provided by the World Bank

Table 7: Leave-One-Out robustness test

Year	Treatment effect	Treatment effect LOO	
		Min	Max
2014	-13.3759	-13.0552	5.2401
2015	-16.3700	-15.8836	4.7669
2016	-17.7069	-17.0131	6.0576
2017	-14.0560	-13.2619	7.9708

Note: The last two columns report the minimum and maximum treatment effect when one control unit with a nonzero weight is excluded at a time.

Source: Data for 1990-2017 provided by the World Bank

To conclude, the placebo tests together with the robustness test, display a weak validity of the results presented in section 6.2. The in-time placebo test points to the fact the treatment effect is somewhat affected by the time of the intervention, but the in-space placebo test together with the robustness test point to the fact that other factors than the intervention itself, especially those affecting Togo, have an impact on the primary school completion rate.

7 Discussion

As presented in section 6, the results presented display weak validity. In accordance with the robustness test, if one of the donor pool countries had been excluded from the sample, the treatment effect of the intervention would have been positive instead of negative. The weak validity of the results subsequently means that the empirical research in this paper is considered to yield a null result, rejecting our hypothesis of a positive relation.

Since the use of the SCM with the selected intervention and research design leads to a null result, one might question if the effect of ICT diffusion on educational attainment is an effect of research design or a signal of an actual zero effect. Previous research displays empirical evidence in support of varying conclusions, both in line with a positive relationship between ICT and education and in line with a zero effect. Empirical support for the positive impact of ICT on educational performance is presented by Kho, Lakdawala, and Nakasone (2018) as well as by Bianchi, Song, and Lu (2020). Bagchi (2005) finds support for a correlation between ICT and educational attainment, controlling for income levels, and Asongu and Odhiambo (2019) find empirical support for a positive impact only for countries with below-median educational quality. In contrast to these studies, several scholars provide empirical evidence for a zero result on the relationship between ICT and educational performance (cf. Faber, Sanchis-Guarner, and Weinhardt 2015; Fairlie and Robinson 2013; Malamud et al. 2018; Bessone, Dahis, and Ho 2020) as well as a zero result on the relationship between ICT and years of schooling (Baliamoune-Lutz 2003). In line with contrasting prior research, the null result yielded in this paper is thus not considered unexpected.

Despite the null result in this particular context, the authors consider the SCM a good identification strategy since the research question is broader than other related studies, where specific mechanisms are isolated. Previous literature has studied particular mechanisms such as assigning computers to children's homes or providing children with laptops using randomised control studies where the causal impact of those specific mechanisms can be isolated (cf. Fairlie and Robinson 2013; Malamud et al. 2018). Since the research question asked in this paper tries to address the national-level effects of broad Internet penetration on primary completion, the chosen methodology provides the means to do so. Furthermore, the comparison between SCM Bias values and AC Bias values in section 6.1.2 prove that the SCM creates a better estimation than an average control of the donor pool, strengthening the argument for the choice of identification strategy.

The treatment effect of above-threshold Internet penetration on primary school completion rate, presented in section 6.2, might be affected by several factors. Firstly, the primary school completion rate of Togo clearly demonstrates that other determinants affect the main outcome variable. Secondly, the application of the SCM to the selected intervention give rise to several methodological limitations that need to be considered in relation to the results.

As presented in Figure 9 (Appendix B), Togo's primary school completion rate has increased substantially from very low rates to the highest completion rate out of donor pool countries and The Gambia in recent years. The substantial improvement in primary school completion rate, despite the fact that Togo is one of the countries with below-threshold Internet penetration, clearly demonstrates that other determinants are important for primary school completion rate. The LOO robustness test conducted in section 6.3.3 demonstrates predicted treatment effects in the case that one of the positively weighted donor pool countries is left out of the sample. The lowest predicted LOO-path shows that the treatment effect of above-threshold Internet penetration would be positive in the case that Togo was excluded from the donor pool. Other determinants thus seem to have a great impact on primary school completion, and this is in line with previous research stating that factors of economic development and religious affiliation impact our main outcome variable (Daun 2000).

The application of SCM in this particular research context furthermore leads to several limitations. The lack of available data, SCM exclusion criteria and the fact that many countries in SSA have reached the critical threshold level for Internet penetration led to a relatively small donor pool with only seven included countries. A larger donor pool might have resulted in a predicted fit more closely resembling the actual pre-intervention outcome. Several countries within the donor pool furthermore lacked data observations leading to the use of mean imputation, although the low percentage of observations replaced using mean imputation should not significantly affect the results. The selection of a critical threshold level for Internet penetration furthermore affects the results as well as the pre-treatment and post-treatment period. The in-time placebo tests (Figure 6 presented in section 6.3.1) demonstrate that the negative treatment effect, albeit with weak validity, exists long before the chosen intervention year. Internet penetration might thus affect educational attainment long before it reaches the threshold of 15 per cent, meaning that the critical threshold might be lower, as suggested by some scholars and as previously mentioned. Lastly, qualitative research was conducted to identify potential idiosyncratic shocks[20], whereas a data-driven approach to identifying such shocks might have improved donor pool validity.

The fact that using the SCM to study the effect of ICT diffusion on education yields a null result in this context does not necessarily mean that ICT has no impact on education. Being able to isolate the causal relationship in this context proved concerning, however, and more research

[20] Since idiosyncratic shocks to the main outcome variable are one exclusion criteria for the SCM.

should be conducted on the topic since scholars reach contrasting conclusions. As suggestions for future research, the authors propose the use of the newly developed Augmented Synthetic Control Method (Ben-Michael, Feller, and Rothstein 2021). The Augmented SCM allows for negative weights, something which could have greatly improved the pre-treatment fit in our paper[21] since many of the donor pool countries have lower primary school completion rates than The Gambia. Furthermore, the research on the relationship between ICT and education should be expanded to differentiate between primary, secondary and tertiary education since ICT might interact differently with education depending on the educational complexity. Lastly, the authors propose broader empirical research defining all determinants of educational attainment in order to understand in which way the interplay of several determinants impacts education.

8 Conclusion

The region of sub-Saharan Africa is particularly affected by digital inequalities and an overrepresentation of children out of school (Roser and Ortiz-Ospina 2013; Roser, Ritchie, and Ortiz-Ospina 2015). While ICTs have been identified by the United Nations as important to achieve SDG 4, related to primary and secondary education, research on the relationship between ICT and education is lacking (Lim et al. 2020). This study addresses the research gap through the application of the SCM and a critical threshold level of Internet penetration to study the impact on primary school completion in SSA. While the evidence for the actual outcome of The Gambia compared with the synthetic counterfactual displays a negative treatment effect on primary school completion, placebo- and robustness tests show weak validity.

Limitations put forth by the SCM in the specific context yield a null result and leave room for future research to address the relationship between ICT and educational outcomes. We encourage further research using the newly developed Augmented SCM, as well as broader research into the interplay of determinants of educational performance and educational attainment. ICT might also interact differently with the school setting depending on the educational level, which is why future studies should expand upon secondary and tertiary education, not covered in this study.

[21] The Augmented SCM developed by Ben-Michael, Feller and Rothstein in 2021 is not yet compatible with STATA to the best of the authors' knowledge, which is why it was not used in this study.

References

Abadie, Alberto. 2021. 'Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects'. *Journal of Economic Literature* 59 (2): 391–425. https://doi.org/10.1257/jel.20191450.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. 'Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program'. *Journal of the American Statistical Association* 105 (490): 493–505. https://doi.org/10.1198/jasa.2009.ap08746.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2015. 'Comparative Politics and the Synthetic Control Method'. *American Journal of Political Science* 59 (2): 495–510. https://doi.org/10.1111/ajps.12116.

Abadie, Alberto, and Javier Gardeazabal. 2003. 'The Economic Costs of Conflict: A Case Study of the Basque Country'. *American Economic Review* 93 (1): 113–32. https://doi.org/10.1257/000282803321455188.

Albiman, Masoud Mohammed, and Zunaidah Sulong. 2016. 'The Role of ICT Use to the Economic Growth in Sub Saharan African Region (SSA)'. *Journal of Science and Technology Policy Management* 7 (3): 306–29. https://doi.org/10.1108/JSTPM-06-2016-0010.

Anděl, J., I. Bičík, and J. D. Bláha. 2018. 'Concepts and Delimitation of the World's Macro-Regions'. *Miscellanea Geographica* 22 (1): 16–21. https://doi.org/10.2478/mgrsd-2018-0001.

Anghel, Brindusa, Antonio Cabrales, Jorge Sainz, and Ismael Sanz. 2015. 'Publicizing the Results of Standardized External Tests: Does It Have an Effect on School Outcomes?' *IZA Journal of European Labor Studies*4 (1): 7. https://doi.org/10.1186/s40174-014-0029-3.

Asongu, Simplice A., and Paul N. Acha-Anyi. 2020. 'Enhancing ICT for Productivity in Sub-Saharan Africa: Thresholds for Complementary Policies'. *African Journal of Science, Technology, Innovation and Development* 12 (7): 831–45. https://doi.org/10.1080/20421338.2020.1732596.

Asongu, Simplice A., and Nicholas M. Odhiambo. 2019. 'Enhancing ICT for Quality Education in Sub-Saharan Africa'. *Education and Information Technologies* 24 (5): 2823–39. https://doi.org/10.1007/s10639-019-09880-9.

Ayanso, Anteneh, Danny I. Cho, and Kaveepan Lertwachara. 2014. 'Information and Communications Technology Development and the Digital Divide: A Global and Regional Assessment'. *Information Technology for Development* 20 (1): 60–77. https://doi.org/10.1080/02681102.2013.797378.

Bagchi, Kallol. 2005. 'Factors Contributing to Global Digital Divide: Some Empirical Results'. *Journal of Global Information Technology Management* 8 (3): 47–65. https://doi.org/10.1080/1097198X.2005.10856402.

Baliamoune-Lutz, Mina. 2003. 'An Analysis of the Determinants and Effects of ICT Diffusion in Developing Countries'. *Information Technology for Development* 10 (3): 151–69. https://doi.org/10.1002/itdj.1590100303.

Ben-Michael, Eli, Avi Feller, and Jesse Rothstein. 2021. 'The Augmented Synthetic Control Method↤'. *Journal of the American Statistical Association* 116 (536): 1789–1803. https://doi.org/10.1080/01621459.2021.1929245.

Bessone, Pedro, Ricardo Dahis, and Lisa Ho. 2020. 'The Impact of 3G Mobile Internet on Educational Outcomes in Brazil'. *Working Paper*, 39.

Bianchi, Nicola, Hong Song, and Yi Lu. 2020. 'The Effect of Computer-Assisted Learning on Students' Long-Term Development', NBER Working Paper Series, no. 28180. https://doi.org/10.3386/w28180.

Bicaba, Zorobabel, Zuzana Brixiová, and Mthuli Ncube. 2017. 'Can Extreme Poverty in Sub-Saharan Africa Be Eliminated by 2030?' *Journal of African Development*19 (2): 93–110.

Bifulco, Robert, Ross Rubenstein, and Hosung Sohn. 2017. 'Using Synthetic Controls to Evaluate the Effect of Unique Interventions: The Case of Say Yes to Education'. *Evaluation Review* 41 (6): 593–619. https://doi.org/10.1177/0193841X17742233.

Cerqua, Augusto, and Giorgio Di Pietro. 2017. 'Natural Disasters and University Enrolment: Evidence from L'Aquila Earthquake'. *Applied Economics* 49 (14): 1440–57. https://doi.org/10.1080/00036846.2016.1218431.

Daun, Holger. 2000. 'Primary Education in Sub-Saharan Africa–a Moral Issue, an Economic Matter, or Both?' *Comparative Education* 36 (1): 37–53. https://doi.org/10.1080/03050060027755.

David, Oladipo Olalekan, and Wynand Grobler. 2020. 'Information and Communication Technology Penetration Level as an Impetus for Economic Growth and Development in Africa'. *Ekonomska Istraživanja* 33 (1): 1394–1418. https://doi.org/10.1080/1331677X.2020.1745661.

Dutta, Amitava, and Rahul Roy. 2005. 'The Mechanics of Internet Growth: A Developing-Country Perspective'. *International Journal of Electronic Commerce* 9 (2): 143–65. https://doi.org/10.1080/10864415.2005.11044329.

Dutta, Soumitra, Irene Mia, World Economic Forum, and INSEAD. 2009. 'The Global Information Technology Report 2008-2009: Mobility in a Networked World'. Geneva: World Economic Forum : INSEAD. http://www.weforum.org/pdf/gitr/2009/gitr09fullreport.pdf.

Faber, Benjamin, Rosa Sanchis-Guarner, and Felix Weinhardt. 2015. 'ICT and Education: Evidence from Student Home Addresses', Working Paper Series, June. https://doi.org/10.3386/w21306.

Fairlie, Robert W., and Jonathan Robinson. 2013. 'Experimental Evidence on the Effects of Home Computers on Academic Achievement among Schoolchildren'. *American Economic Journal: Applied Economics* 5 (3): 211–40. https://doi.org/10.1257/app.5.3.211.

Gillwald, Alison, and Onkokame Mothobi. 2019. 'After Access 2018: A DEMAND-SIDE VIEW OF MOBILE INTERNET FROM 10 AFRICAN COUNTRIES'. Policy Paper Series No.5 After Access Paper No. 7 (2018). Research ICT Africa.

Government of The Gambia. 2017. 'The Gambia National Development Plan (2018-2021)'. https://mofea.gm/downloads-file/national-development-plan.

Hjort, Jonas, and Lin Tian. 2021. 'The Economic Impact of Internet Connectivity in Developing Countries'. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3964618.

Kho, Kevin, Leah Lakdawala, and Eduardo Nakasone. 2018. 'Impact of Internet Access on Student Learning in Peruvian Schools'. *Working Papers*, April. https://ideas.repec.org/p/ris/msuecw/2018_003.html.

Lechman, Ewa. 2015. *ICT Diffusion in Developing Countries*. https://doi.org/10.1007/978-3-319-18254-4.

Leroutier, Marion. 2022. 'Carbon Pricing and Power Sector Decarbonization: Evidence from the UK'. *Journal of Environmental Economics and Management* 111 (January): 102580. https://doi.org/10.1016/j.jeem.2021.102580.

Lewin, Keith M. 2009. 'Access to Education in Sub‐Saharan Africa: Patterns, Problems and Possibilities'. *Comparative Education* 45 (2): 151–74. https://doi.org/10.1080/03050060902920518.

Lim, Cher Ping, Sungsup Ra, Brian Chin, and Tianchong Wang. 2020. 'Information and Communication Technologies (ICT) for Access to Quality Education in the Global South: A Case Study of Sri Lanka'. *Education and Information Technologies* 25 (4): 2447–62. https://doi.org/10.1007/s10639-019-10069-3.

Malamud, Ofer, Santiago Cueto, Julian Cristia, and Diether W. Beuermann. 2018. 'Do Children Benefit from Internet Access? Experimental Evidence from Peru', Working Paper Series, November. https://doi.org/10.3386/w25312.

McClelland, Robert, and Sarah Gault. 2017. 'The Synthetic Control Method as a Tool to Understand State Policy'. Urban Institute.

Niebel, Thomas. 2014. 'ICT and Economic Growth - Comparing Developing, Emerging and Developed Countries'. *ICT and Economic Growth*, no. 14: 29.

Pieters, Hannah, Daniele Curzi, Alessandro Olper, and Jo Swinnen. 2014. 'Political Reforms and Food Security'. https://doi.org/10.22004/ag.econ.173092.

Porter, Gina, Kate Hampshire, James Milner, Alister Munthali, Elsbeth Robson, Ariane Lannoy, Andisiwe Bango, et al. 2016. 'Mobile Phones and Education in Sub-Saharan Africa: From Youth Practice to Public Policy'. *Journal of International Development* 28 (1): 22–39. https://doi.org/10.1002/jid.3116.

Rainie, Lee, and Barry Wellman. 2019. 'The Internet in Daily Life: The Turn to Networked Individualism'. In *Society and the Internet*, 2nd ed. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198843498.003.0002.

Roser, Max. 2013. 'Economic Growth - Our World in Data'. Our World in Data. https://ourworldindata.org/economic-growth.

Roser, Max, and Esteban Ortiz-Ospina. 2013. 'Primary and Secondary Education - Our World in Data'. Our World in Data. 17 July 2013. https://ourworldindata.org/primary-and-secondary-education.

Roser, Max, Hannah Ritchie, and Esteban Ortiz-Ospina. 2015. 'Internet - Our World in Data'. Our World in Data. 14 July 2015. https://ourworldindata.org/internet.

Stanley, T. D., Hristos Doucouliagos, and Piers Steel. 2018. 'DOES ICT GENERATE ECONOMIC GROWTH? A META-REGRESSION ANALYSIS'. *Journal of Economic Surveys* 32 (3): 705–26. https://doi.org/10.1111/joes.12211.

'The Gambia Education Sector Strategic Plan 2016-2030'. 2017. Ministries of Basic and Secondary Education and Higher Education Research Science and Technology. https://www.globalpartnership.org/content/2016-30-education-sector-plan-gambia

UNDP. n.d. 'About Africa'. UNDP in Africa. Accessed May 15, 2022. https://www.africa.undp.org/content/rba/en/home/regioninfo.html.

UNESCO. 2021. 'Information and Communication Technology (ICT) in Education | Unesco IIEP Learning Portal'. UNESCO's International Institute for Educational Planning. 13 July 2021. https://learningportal.iiep.unesco.org/en/issue-briefs/improve-learning/information-and-communication-technology-ict-in-education.

United Nations. n.d. 'Goal 4 | Department of Economic and Social Affairs'. Accessed May 15, 2022. https://sdgs.un.org/goals/goal4.

United States Institute of Peace. n.d. 'Independence of South Sudan'. United States Institute of Peace. Accessed May 15, 2022. https://www.usip.org/programs/independence-south-sudan.W

Verwimp, Philip, and Jan Van Bavel. 2014. 'Schooling, Violent Conflict, and Gender in Burundi'. *The World Bank Economic Review* 28 (2): 384–411.

Wamboye, Evelyn, Kiril Tochkov, and Bruno S. Sergi. 2015. 'Technology Adoption and Growth in Sub-Saharan African Countries'. *Comparative Economic Studies* 57 (1): 136–67. http://dx.doi.org/10.1057/ces.2014.38.

World Bank. n.d.a. 'Individuals Using the Internet (% of Population) - Sub-Saharan Africa, European Union | Data'. Accessed May 15, 2022. https://data.worldbank.org/indicator/IT.NET.USER.ZS?locations=ZG-EU&view=chart.

World Bank. n.d.b. 'Primary Completion Rate, Total (% of Relevant Age Group) | Data'. Accessed May 15, 2022. https://data.worldbank.org/indicator/SE.PRM.CMPT.ZS.

World Bank. n.d.c. 'Sub-Saharan Africa | Data'. Accessed May 15, 2022. https://data.worldbank.org/country/ZG.

World Bank. n.d.d. 'The Gambia - Overview'. Accessed May 15, 2022. https://www.worldbank.org/en/country/gambia/overview.

Appendices

Appendix A: Data

Table 8: Data information

Variable name	Descrpition	Data period	Source	Code
Primary school completion rate	Primary completion rate, total (% of relevant age group)	1990-2017	World Bank Open Data	SE.PRM.CMPT.ZS
Pupil-teacher ratio	Pupil-teacher ratio, primary	1990-2017	World Bank Open Data	SE.PRM.ENRL.TC.ZS
Probability of children dying 5 to 9	Probability of dying among children ages 5-9 years (per 1,000)	1990-2017	World Bank Open Data	SH.DYN.0509
Probability of children dying 10 to 14	Probability of dying among children ages 10-14 years (per 1,000)	1990-2017	World Bank Open Data	SH.DYN.1014
Population ages 0 to 14	Population ages 0-14 (% of total population)	1990-2017	World Bank Open Data	SP.POP.0014.TO.ZS
GDP PC constant 2010	GDP per capita (constant 2010 US$)	1990-1991, 2012-2017	International Monetary Fund (IMF)	NGDP_RPCH
GDP PC constant 2010	GDP per capita (constant 2010 US$)	1992-2011	World Bank Open Data	NY.GDP.PCAP.KD
Individual using Internet	Individuals using the Internet (% of population)	1990-2017	World Bank Open Data	IT.NET.USER.ZS

Note: Data was collected February 10th 2022

Appendix B: Additional Results

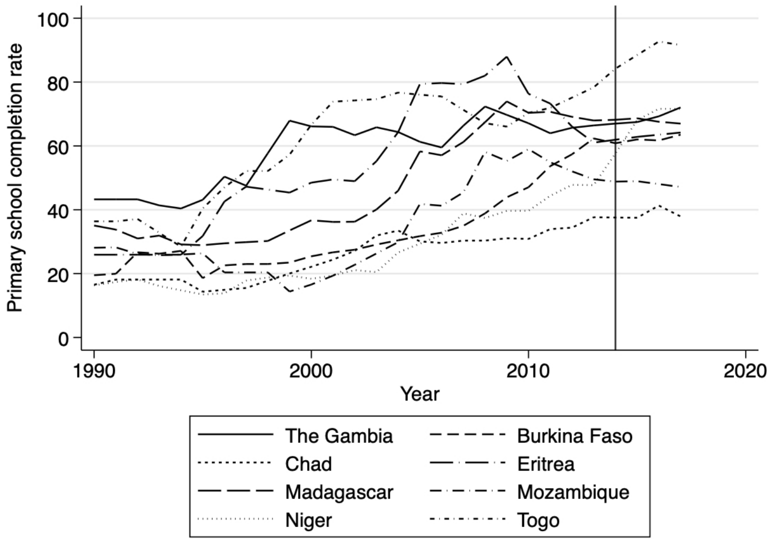

Figure 9: Primary school completion rate

Source: Data for 1990-2017 provided by the World Bank

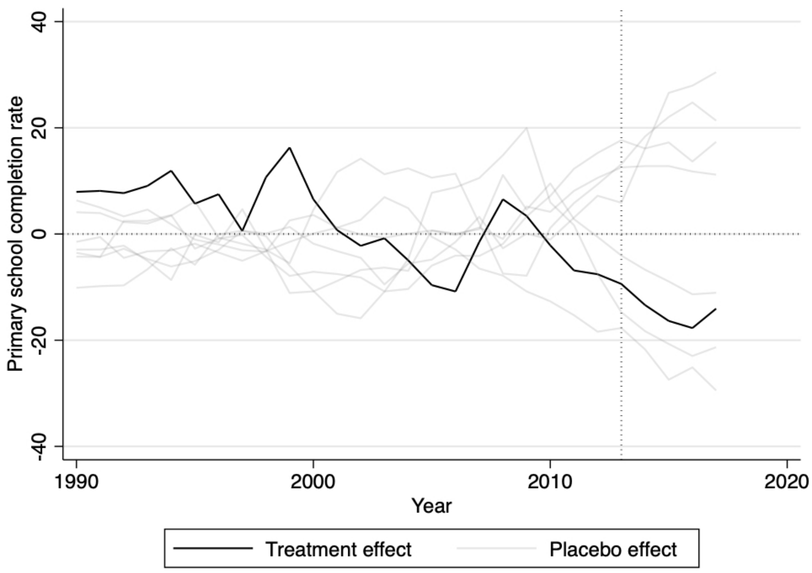

Figure 10: In-space placebo test excluding countries with above two times RMSPE

Source: Data for 1990-2017 provided by the World Bank