HETEROGENEOUS IMPACTS OF SCHOOL FEE ELIMINATION IN TANZANIA: GENDER AND COLONIAL INFRASTRUCTURE

Roxana MANEA, Pedro NASO

Heterogeneous Impacts of School Fee Elimination in Tanzania: Gender and Colonial Infrastructure

Roxana Manea[*] and Pedro Naso[†]

May 8, 2021

Abstract

In this study, we investigate the impacts of the 2002 elimination of primary school fees in Mainland Tanzania. We explore how the magnitude of these effects depends on gender and the size of early investments in the educational infrastructure of Tanganyika. We use the 2002 and 2012 census waves as well as historical information on the location of schools in the late 1940s, and conduct a difference-in-differences analysis. We find that exposure to an average of 1.7 years of free primary education has reduced the proportion of people who have never attended primary education by 6.8 percentage points. The benefits of fee removal have been significantly larger for females compared to males, and females from districts where the size of investments in education was relatively larger during colonial rule have been the greatest beneficiaries.

JEL classification: I28, J16, N37.

1 Introduction

The evidence regarding the persistence of historical events and their impacts on current socio-economic outcomes is well documented (Nunn, 2009, 2014b). It is argued that culture (Nunn & Wantchekon, 2011; Alesina et al., 2013), the existence of multiple equilibria (Redding et al., 2011; Bleakley & Lin, 2012) and the shaping of domestic institutions (Acemoglu et al., 2005; Dell, 2010; Vogler, 2019) facilitate the persistence of the impact of historical events. In contrast, evidence regarding how the persistence of historical institutions and events affect the outcomes of contemporary reforms is limited. Governments implement reforms with the objective of improving outcomes on average as well as to even out the distribution of outcome values across individuals. Disparities exist for a variety of reasons. For instance, historical institutions and infrastructure have been linked to the uneven distribution of resources and socio-economic outcomes—whether by favouring one group over the general population (Becker & Woessmann, 2008; Huillery, 2009; Calvi & Mantovanelli, 2018) or by holding back some populations (Frankema, 2010; Bruhn & Gallego, 2012; Naritomi et al., 2012). Thus, we ask these questions: How do current reforms interact with the legacies of historical events and institutions? Do they mitigate or perpetuate historical disparities?

We show that current development policies may feed the disparities created by history as opposed to eliminating them. We evaluate the impacts of the 2002 elimination of primary school fees in Tanzania. Then, we explore the interaction of these impacts with the early colonial and missionary infrastructure of Tanganyika, which has created a pattern whereby the gender gap in education is smaller in districts that have benefited from relatively larger investments during colonial rule.

We find that the elimination of fees has improved educational outcomes on average. We estimate a 6.8 percentage point reduction in the ratio of people who have never enrolled in school and a 6-month improvement in time spent in primary education.[1,2] Moreover, the reform has also reduced the educational gender gap. Furthermore, females from districts with a stronger history of investments in education have been the greatest beneficiaries of the reform even though they were already better off compared to other females. For instance, a standard deviation increase in the number of Protestant or Catholic schools scaled per 100,000 students leads to a further reduction of 1.7 percentage points in the ratio of females who have never enrolled in primary education. Females from districts with historically low investments in education would have required the largest boost. Nevertheless, they were held back as their female counterparts from districts with denser school coverage in the 1940s benefited disproportionately. Finally, we compile evidence that one of the mechanisms at play is the persistence of investments. Districts with a larger educational infrastructure in the past continue to have a larger infrastructure today.

---

[1]Estimated for an average reform intensity of 1.7 additional years of free education.

[2]These effects apply to cohorts that were of school age at the time of the reform or shortly after. The persistence of these impacts should be studied again when more recent data becomes available. We caution that this is not the first time Tanzania removed fees in a bid to raise enrolment rates (World Bank & UNICEF, 2009; Somerset, 2009).

We employ a difference-in-differences estimation strategy with multiple layers of fixed effects. We use a panel of gender-age-district cohorts observed in 2002 and 2012. This panel is merged with historical data on the educational infrastructure of Tanganyika around the late 1940s. For this purpose, we have geo-referenced a map documented by Buchert (1991), which includes information on the type of school infrastructure: Catholic, Protestant or village authority. To estimate the magnitude and heterogeneity of reform impacts, we had to overcome a serious obstacle posed by the simultaneous, country-wide implementation of the reform. We have exploited the fact that some individuals are treated while others are not, depending on one's year of birth. Additionally, for those who are treated, the intensity of their exposure to the reform depends on the pre-reform educational performance of their district. For instance, in districts with low pre-reform performance, exposure to the benefits of free primary education will be at its highest. This identification strategy has already been used to evaluate the impacts of fee elimination in Ethiopia (Chicoine, 2019, 2020), Kenya (Lucas & Mbiti, 2012a,b) and Malawi (Zenebe Gebre, 2019). We elicit reform heterogeneities by interacting the reform with variables denoting gender and the colonial school infrastructure of Tanganyika. The parallel trends assumption is validated and strongly supportive of our difference-in-differences strategy.

The contribution of our study is threefold. First, we bring additional evidence to the literature discussing the implementation of free primary education. Our results are similar to those estimated for other African countries, although the magnitude of our estimates is slightly smaller than in the case of Ethiopia (Chicoine, 2019) and less than half of the effects found by Zenebe Gebre (2019) for Malawi. Deininger (2003), Nishimura et al. (2008) and Grogan (2009) also confirm our findings. The authors evaluate the 1997 fee reform of Uganda and show that girls have experienced larger improvements relative to boys. In contrast, evaluations of the Kenyan reform find a higher impact on the graduation rates of boys compared to girls, thus contributing to a widening of the gender gap in favour of boys (Lucas & Mbiti, 2012b). Regarding evidence on Tanzania, Hoogeveen & Rossi (2013) have already attempted to evaluate the elimination of fees. They exploit the staggered absorption of out-of-school children as an identification strategy. In line with our own findings, they show that the elimination of fees has increased the probability of children being in school at age 7. Unlike our study, they are unable to assess the effects of the reform on completed education due to data limitations. We further refine their findings by using a different identification methodology and exploring the heterogeneity of impacts with respect to the historical infrastructure of Tanzania.

Second, we contribute to the literature assessing the long-term impacts of historical events on education. The study closest to ours is the cross-country analysis of Nunn (2014a). Similar to his findings, we argue that the investments of Protestant missions have had positive impacts on the educational outcomes of females relative to males. Becker & Woessmann (2008) have also put forward this hypothesis. However, unlike Nunn (2014a) who does not find any significant effects of Catholic schools on female education, we do find evidence that these schools are associated with smaller educational gaps between females and males in Tanzania. Finally, Montgomery (2017) also evaluates the impacts of the educational infrastructure erected during colonial rule in Tanzania. Whereas Montgomery

(2017) argues that the infrastructure of German East Africa is linked to higher gender-based educational gaps in present-day Tanzania, we find that the British and missionary infrastructure of the 1940s is associated with smaller gaps in the long run.

Third and last, we contribute to a relatively limited strand of the literature which shows that historical legacies inform the success of reforms down the line (Pop-Eleches, 2007). Similarly, the patterns created by other phenomena that are characterised by persistence, i.e., traditions, have also been shown to endure in the aftermath of reforms (Ashraf et al., 2020). For instance, girls raised in the bride price tradition are better educated, and educational policies have benefited them more relative to girls from other traditions. The hypothesised channel is the marriage market (Ashraf et al., 2020). In this paper, we show that Tanganyika's historical infrastructure speaks to the distribution of current reform impacts because of the spatial persistence of investments in educational infrastructure.

2 Context

2.1 Educational Infrastructure during Colonial Rule

Modern-day Mainland Tanzania was German East Africa between 1891 and the end of World War I. Thereafter, Britain gained control over the area, which became known as Tanganyika. Although present since the 1840s, mission schools were initially few and far between (Buchert, 1991). Teaching was conducted in local languages (Tabetah, 1982), and the main focus was to "*civilise*" and convert the local populations to Christianity (Tabetah, 1982; Buchert, 1991). Mission schools expanded geographically during the German colonial rule (Gillette, 1977; Buchert, 1991), and while proselytisation remained their most important goal, they also incorporated secular teachings (Buchert, 1991). Mission schools were open to both boys and girls (Gillette, 1977).

Besides the missionary infrastructure, the government also set up its own schools. Its concerns were non-religious and focused on vocational, civic and general education (Gillette, 1977). However, the government infrastructure was limited, clustered around the coast and only served the male children of chiefs (Gillette, 1977). As of 1913, there were roughly 115,000 children in school and 95 percent of them were educated in missionary schools (Buchert, 1991). Nevertheless, most of this educational groundwork was substantially damaged during World War I (Cameron & Dodd, 1970; Chachage, 1988; Buchert, 1991).

During the interwar period, the British administration promoted a laissez-faire philosophy and avoided assuming any active role in the educational sector (Cameron, 1967; Buchert, 1991). The colonial state adopted policies that favoured adaptation so as to "*preserve traditional societies*" (Buchert, 1991). This translated into a vocational curriculum, which relied on the use of local languages and the implementation of agricultural and practical activities, with the ultimate purpose of preparing and training Africans to accept and serve the economic and political goals of the colony (Buchert, 1991).

This policy of non-interference applied to government-assisted schools; however, these were in limited supply during the interwar period (Cameron, 1967). The overwhelming majority of the educational infrastructure was still made up of mission schools that were *not* assisted by the state and which chiefly brought literacy skills to converts as a result of religious teachings (Buchert, 1991). In 1924, 21 percent of children were in school: 5,000 attended 72 public schools and 162,000 children studied in mission schools (Siwale & Sefu, 1977).

As a result of Britain's system of indirect rule, which relied on local leaders to govern the territory, revenues from taxation provided traditional authorities with an independent source of finance to develop their local infrastructure, e.g., schools (Chachage, 1988; Buchert, 1991). Consequently, village- or "*native-*" authority schools, were set up in addition to government and mission schools. The local-authority schools had similar aims as government schools. They did not have any religious goals but focused on educating local chiefs, headmen and their children (Cameron & Dodd, 1970). Ultimately, these schools came under government control (Gillette, 1977), although there were a few exceptions, such as the schools set up by the Chagga people who continued to use revenues from their cash crops to fund and supervise their schools (Cameron & Dodd, 1970).

World War II spared Tanganyika in ways that World War I did not. World War II caused a rise in the prices of primary goods, some of which were produced by Tanganyika. Thus, state revenues increased (Cameron & Dodd, 1970). Moreover, the role of education started to become widely recognised, as the colonial state required skilled labour to meet the needs of a modern economy which relied on the production of cash crops (Buchert, 1991). Furthermore, international pressure for social development mounted (Cameron, 1967; Siwale & Sefu, 1977). Consequently, the British colonial state adopted a new, interventionist philosophy as a provider of financial resources and inspector of school activities (Cameron, 1967; Cameron & Dodd, 1970; Buchert, 1991). The active involvement of the British administration in the educational sector has ultimately favoured the country-wide progressive migration from a vocational to an academic curriculum. Mission schools had to conform to the centrally-determined curriculum, which reduced the importance of religious teachings, so as to be eligible for financial assistance (Cameron, 1967; Buchert, 1991). In fact, unassisted mission schools saw their numbers decrease (Buchert, 1991). In 1945, they represented roughly 60 percent of all mission schools (Cameron & Dodd, 1970). Instruction in local languages was virtually replaced by Swahili in all types of schools. The basic, four-year primary school cycle included classes on arithmetic, reading, writing, religion, health and hygiene, general knowledge, physical education, singing, agriculture and handwork (Cameron & Dodd, 1970).

Throughout colonial rule, school teachings have been gender-specific regardless of school type, with girls usually being disadvantaged both in quantitative as well as qualitative terms (Mbilinyi et al., 1991; Olekambaine, 1991). At the same time, however, the colonial state was not particularly invested in the education of African boys either. The gender ratio was roughly 1 girl to 3 boys among primary school students in 1956 (Gillette, 1977; Siwale & Sefu, 1977). Missionaries favoured the numbers, "*the production of sincere, edu-*

cated Christians, of whom the more the merrier" and only offered rudimentary education (Cameron & Dodd, 1970), "*enough to understand the Bible but not so much as to result in pupils turning away from the church*" (Gillette, 1977). In contrast, government schools were fewer, had fewer students as well, and were more preoccupied with the creation of individuals ready to take on administrative tasks in the government apparatus.

2.2 Post-independence Educational Developments

In 1958, three years before independence, only 24 percent of children aged 5–14 were in school in Mainland Tanzania (Omari et al., 1983). Despite this situation, the post-independence government was slow to expand the primary education sector during the 1960s (Omari et al., 1983; Oketch & Rolleston, 2007). As of the late 1960s, girls continued to be under-represented among primary school students. In 1968, they represented 39 percent of an enrolled population of 750,000 children (Gillette, 1977; Siwale & Sefu, 1977).

The Arusha Declaration (1967) was the first step toward the concept of Education for All in a bid to build a Tanzanian national identity (Jerve, 2006). Then, the Musoma Resolution of 1974 declared Universal Primary Education a national priority. Other reforms included the first attempt to remove primary school fees in 1973-74 (Galabawa, 1990; World Bank & UNICEF, 2009), the building of schools[3] and the introduction of compulsory schooling laws in 1978. The educational reforms of the 1970s had finally allowed schools to reach full parity around 1984–85 (Olekambaine, 1991). The net and gross primary school enrolment rates in 1981 were 69.7 and 98.3 percent, respectively (Ishumi, 2014).

The educational outcomes of the early 1980s were encouraging; however, the expansionary efforts of the government also exacted a toll on the quality side of education. To meet demand and stay within budget, most of the newly deployed teachers were insufficiently qualified (Omari et al., 1983; Galabawa, 1990). Despite their increasing numbers, classrooms were also often insufficient to absorb new entrants. Lastly, because of high population growth rates and the adverse economic conditions of the 1980s[4], the quality of the educational system deteriorated further (Galabawa, 1990; Jerve, 2006). Fees were reinstated and education-related costs increased (Jerve, 2006). By 1990, the net and gross enrolment rates had decreased to 59.6 and 80.7 percent, respectively (Ishumi, 2014).

2.3 Free Primary Education

Following the deteriorating educational situation of the late 1980s and 1990s, the Government of Tanzania adopted the Primary Education Development Plan (PEDP) (Government of Tanzania, 2001). Per this plan, primary school fees and all other mandatory

---

[3]There were 3,238 primary schools in 1961, 4,070 in 1970 and 9,931 in 1980 (Ishumi, 2014).

[4]These adverse conditions amounted to the oil crises of the 1970s, Tanzania's war with Uganda and agricultural stagnation (Buchert, 1991; Jerve, 2006; Vavrus & Moshi, 2009).

parental contributions were removed as of January 2002.[5,6] To compensate schools for the loss in income, the PEDP introduced a capitation grant of US \$10 (TSh9,000) per child per year as well as an investment grant to build the necessary classrooms, sanitation facilities and teachers' accommodation. The capitation grant was increased to TSh10,000 in 2006 (Government of Tanzania, 2006) and has since remained constant (Mbiti et al., 2019).

Table 1: Pre- and Post-reform Infrastructure

		PRE-REFORM		POST-REFORM	
		1999	2001	2002	2005
(1)	Population of 7–13 children[a]	5,427,156	5,679,676	5,810,309	6,220,512
(2)	Gross enrolment rate	77.2%[b]	84%[c]	99%[c]	105.41%[d]
(3) = (1) × (2)	Children in school	4,189,764	4,770,928	5,752,206	6,557,042
(4)	Stock of classrooms start of year	57,367[e]	est. 60,000[e]	est. 60,000[e]	89,875[f]
(5)	Stock of teachers start of year	103,966[g]	102,313[h]	est. 109,665[i]	134,638[j]
(6) = (3) ÷ (4)	Student-to-classroom ratio	73:1	80:1	96:1	73:1
(7) = (3) ÷ (5)	Student-to-teacher ratio	40:1	47:1	52:1	49:1

[a]World Bank (2001), Annex H, Table 3 and PEDP, Annex 3, Table 1. [b]World Bank (2001), Annex H, Table 3. [c]World Bank (2005), Chapter 4, Table 1. [d]World Bank DataBank and UNESCO Institute for Statistics, ID: SE.PRM.ENRR. [e]World Bank (2001), Annex H, Table 5. [f]Stock of 1999 plus the project-declared output of 29,922 classrooms and plus the output of 2,586 classrooms built under a related World Bank Project (World Bank, 2005). [g]World Bank (2005), Annex H, Table 4. [h]Based on the 1999 stock less attrition at 1.59%. Teacher hirings were frozen (World Bank, 2001). [i]Considering the 1999 stock and attrition rates of approx. 1.59% per year between 1999 and 2000, the stock would have been of 100,665 at the end of 2001 (World Bank, 2001). However, the government planned to hire approx. 9,000 teachers by 2002. This stock of unemployed teachers is likely to have existed because of a prior freeze on teacher hirings (World Bank, 2001). The target would have been missed only if deployment had failed. In a bid not to underestimate the 2002 teacher capacity, we assume all 9,000 teachers were recruited. [j]Stock of 2001 plus the project-declared output of 32,325 teachers (World Bank, 2005). The strategy of temporarily employing double-shift teaching has underperformed. (World Bank, 2005).

The expected surge in enrolment due to the elimination of fees has motivated the government to adopt several measures to avoid overwhelming the educational system, e.g., implementing a staggered absorption of children and expanding the educational infrastructure. The government built 29,922 classrooms and hired 32,325 teachers during 2002–04. Although impressive, these efforts barely managed to maintain educational services at pre-reform levels. See Table 1. While the increase in enrolments at the start of 2002 was notable and a considerable break from previous trends, the change in infrastructure was minimal in 2002. Relative to the number of students, the infrastructure shrank.

---

[5]Zanzibar is an autonomous administrative region and has followed a different reform schedule. The timing of the census data does not allow for the study of the reform in Zanzibar.

[6]Before their elimination, school fees were estimated to have been roughly US \$4.6 per child per academic year (Valente, 2019). Monthly food and non-food expenditure per capita was estimated at TSh10,120 (US \$12.5) for Mainland Tanzania in 2000–01 (National Bureau of Statistics, 2002). This included the monetary equivalent of the food grown by the household. Moreover, Sumra (2017) documents that the overall education-related costs, fees and parental contributions included, were US \$8 to 16 per year per child—the equivalent of one to two months' worth of agricultural wages.

3   Data and Summary Statistics

We use two sources of data to estimate the impact of the elimination of fees on educational outcomes. First, we rely on the 2002 and 2012 Tanzanian census data which were made available by the Tanzanian Bureau of Statistics and distributed by the Minnesota Population Center (2018). Second, we have geo-referenced the map in Figure 1, which was originally included in a 1947 report by His Majesty's Government on the administration of Tanganyika and documented by Buchert (1991).

3.1   Census Data

The census data are collapsed such that the units of observation consist of gender-age-district-year groups. The administrative borders of districts in Tanzania have changed between census years. In order to present the 2002 and 2012 data in terms of the same administrative demarcations, we use the 1988 district borders, as both the 2002 and 2012 administrative units can be traced back to their more encompassing 1988 polygons. Consequently, we have a panel where gender-age-district groups are observed twice.

Table 2: Summary Statistics and T-tests: Census Data

	Mean	σ	Min	Max	Mean Fem.	Mean Males	Diff.	t-stat	p-val
Ratio without any education	0.164	0.123	0.000	0.816	0.188	0.141	0.047	16.33	0.00
Years primary educ. (all individuals)	5.471	0.939	1.236	6.979	5.358	5.585	-0.227	-10.19	0.00
Years primary educ. (only enrollees)	6.530	0.360	4.091	6.989	6.577	6.482	0.095	11.11	0.00
Gender-age-district-period groups	7,004	7,004	7,004	7,004	3,502	3,502	-	-	-

Sample weights have been used to compute all variables. 103 Mainland districts as per the 1988 demarcations $\times$ 2 genders $\times$ 2 periods (2002 and 2012) $\times$ 17 age groups (14 to 30 years old) gives 7,004 observations. For the t-test, $H1$ is Difference $\neq 0$.

The summary statistics presented in Table 2 show that there is a statistically significant difference between women and men in terms of their school attendance, with 19 percent of women never having attended school in the 2002 and 2012 census waves, while only 14 percent of men experienced the same situation. Furthermore, once in school, females seem to acquire more years of education than their male counterparts do. The difference is slight but significant. However, because girls are more likely to be out of school than boys, the overall girl-boy educational gap is significant, with girls performing worse.

3.2   Colonial Infrastructure Data

Map 1 presents the educational school infrastructure of Tanganyika around the late 1940s. We have geo-referenced 1,291 schools of which 177 are government/village-authority schools, 634 are Catholic mission schools and 480 are registered as Protestant. The map suggests that while there are denominational clusters, e.g., Catholic schools in and around the Ruvuma region and Protestant schools in Kilimanjaro, there are also areas with a fair amount of mixing, such as Kagera. Government schools accompany mission schools, but

they also cover remote areas which were otherwise poorly served by mission schools. The map includes additional information that we do not use. For instance, we exclude secondary and teacher training schools from the analysis as well as the schools for European and Asian children. They represent a very small fraction of the total number of schools.

Figure 1: Educational Facilities in Tanganyika Cca. 1947
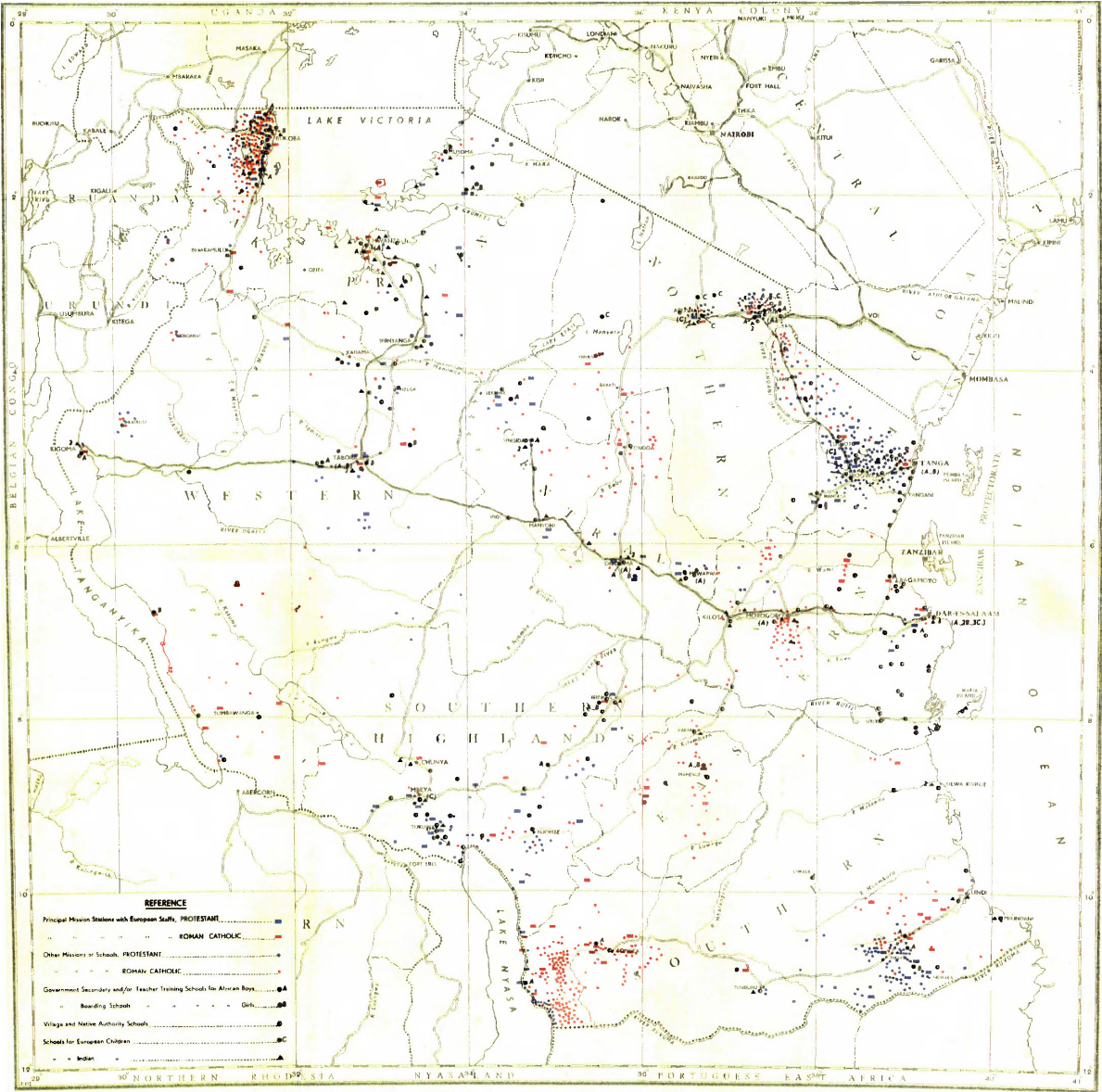
Source: Report by His Majesty's Government on the administration of Tanganyika, 1947. Documented by Buchert (1991), PhD Dissertation. International and Comparative Education, University of London.

In order to account for the fact that population clusters will be accompanied by a larger number of schools, we scale the number of schools to district population. For lack of data regarding population numbers in the 1940s, we use the 2012 population of individuals aged 7–13. The assumption is that the distribution of the 1940s population among districts

is correlated with that of 2012.[7] Huillery (2009) also uses population numbers to scale historical institutions, but she does have access to historical population records. We opine that the use of populations as opposed to surface, i.e., Nunn (2014a), is advised for the scaling of variables in the case of Tanzania, as large swaths of land are uninhabited and make districts appear large without cause. Table 3 shows that an average district in 1947 was endowed with 17 schools per 100,000 children aged 7–13 in 2012. At the same time, some districts had no access to schools, while some had upwards of 100 schools.

Table 3: Summary Statistics: Colonial Infrastructure

Variables	Mean Sample	%	σ	Min	Max
All	17	100	23	0	136
Catholic	8	47	16	0	115
Protestant	6	35	12	0	86
Village authority	3	18	4	0	31
Districts	103	103	103	103	103

Number of schools scaled per 100,000 individuals aged 7–13 in 2012.

3.3   Gender Gap and Colonial Infrastructure

In Figure 2, we map the intensity of the colonial school infrastructure. The rectangle shapes denote the gap in enrolment between males and females. Green is for male advantage and red is for female advantage. The extent to which a rectangle is filled is indicative of the magnitude of the gap. The min and max values of the rectangle shapes are all the same across the maps to allow comparisons, i.e., 0 and 17 percent. The black horizontal lines that cut the rectangle shapes are the mean values for the control and treated groups, computed separately.

Several insights emerge from Figure 2. Enrolment gaps are substantially larger among the control cohorts than they are among the treated cohorts. We also notice a clustering of larger enrolment gaps in areas that had a lower density of schools at the end of the 1940s. We further note that for the case of districts with a denser school infrastructure in the 1940s, such as those marked with an orange triangle, the gap was substantially reduced for the treated cohorts. In contrast, for districts with low infrastructure, such as those marked with an orange circle, the mitigation of the enrolment gap was relatively modest. In the following sections, we explore these relationships using a regression analysis based on a two-period panel of gender-age-district cohorts.

---

[7]By consulting Figure 1b on page 141 of Trewartha & Zelinsky (1954) and Map 2 on page 7 of the Report on the 2012 Population and Housing Census (National Bureau of Statistics & Office of Chief Government Statistician, 2013), we acquire suggestive evidence that the population clusters of the early 1950s have largely stayed in place and overlap with those of 2012.

Figure 2: Never Enrolled Female to Male Differences and Colonial Infrastructure

Control

Scaled 1947 schools
- 20.7 – 136.4
- 8.1 – 20.6
- 4.5 – 8
- 0 – 4.4

Treated
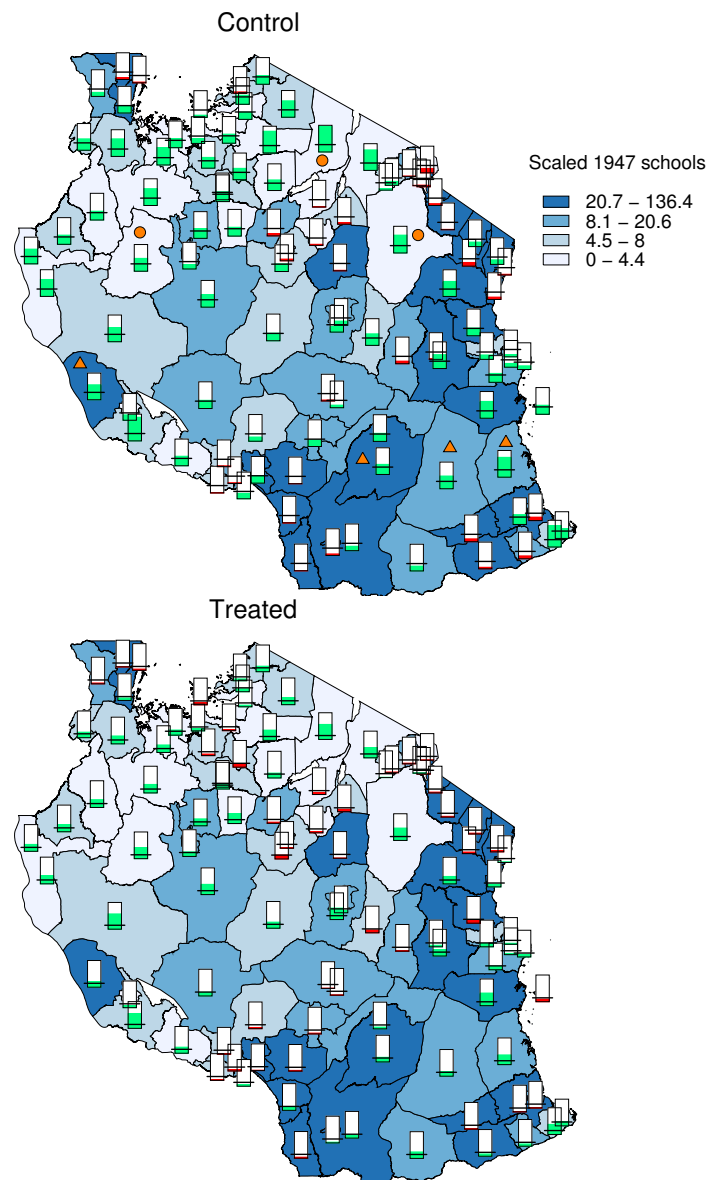
The *rectangular diagrams* plot the difference between the ratio of females and males who have never enrolled in school. Green means a male advantage. Red is for female advantage. The *bottom and top lines of the rectangular* diagrams are the min and the max of the gap taken across both years. Therefore, the gaps plotted on the two maps are comparable. The *horizontal, black line* which crosses the rectangle shapes is the mean computed for the control and treated groups, respectively. Only the 14–23 age cohorts are considered. In 2002, they were not yet treated, but in 2012, they are all either partially or fully treated.

4   Identification Strategy

4.1   Reform Package

The elimination of school fees was part of a wider package of interventions that were put forward by the PEDP (Government of Tanzania, 2001), among which were the building of classrooms and the hiring of teachers. Consequently, we risk confounding the impact of school fee elimination with that of the infrastructure development. However, due to the

timing mismatch between the enrolment surge, the construction of classrooms and the hiring of teachers, which are documented in Table 1, it becomes apparent that a larger number of parents decided to send their children to school as a result of the removal of fees rather than the infrastructure developments. The jump in enrolments was immediate, as was the elimination of fees, while infrastructure developments took time to materialise.

Therefore, we argue that it was the removal of fees rather than these infrastructure developments which paved the way for improvements in education post-reform. Hoogeveen & Rossi (2013) adopt the same approach for their case study on Tanzania, and so do Deininger (2003), Grogan (2009) and Nishimura et al. (2008) in their studies of the Ugandan fee reform. Per this argument, infrastructure improvements have chiefly enabled the absorption of new enrolments, which were in return driven by the removal of fees.

4.2   Reform Intensity

An additional identification problem is posed by the fact that school fees were eliminated simultaneously across the territory of Mainland Tanzania. This complicates our identification mission because the reform may have overlapped with other country-wide developments which risk confounding the impacts of the reform. To address this issue, we employ a methodology that was initially used to evaluate the impacts of anti-malaria interventions on education and fertility in Sri Lanka and Paraguay (Lucas, 2010, 2013). Ultimately, this methodology was also applied to evaluate educational reforms that were implemented country-wide at the same time (Lucas & Mbiti, 2012a,b; Chicoine, 2019, 2020). The strategy relies on the fact that some individuals are treated while others are not, depending on one's year of birth and consequent age at the time when fees were eliminated. Moreover, for those who are treated, their potential response to the reform depends on their district's pre-reform educational performance. That is, the reform has had various degrees of geographical intensity. Districts which were performing poorly pre-reform will have a higher potential to improve following the elimination of school fees, while those with already satisfactory performance will have relatively less room for improvement, which means the intensity of the reform will be lower for the latter districts.

The reform intensity variable takes value zero for all age groups that were 14 or older in 2002 (born in or before 1988) and non-zero increasing values for cohorts aged 8–13 in 2002 (born between 1989 and 1994), as they were in school already when primary school fees were eliminated. Seven-year-olds and younger age groups (born in or after 1995) have been fully exposed to the reform and are assigned the highest values of reform intensity.[8] At the same time, intensity varies across districts and genders, as pre-reform educational attainment is averaged for each gender in each district. For the benchmark indicator, the pre-reform gender-district educational attainment is computed using the 2002 census, namely data from 14 cohorts born between 1970 and 1983, 19–32 years old in 2002. As a robustness check, we have also computed the intensity indicator based on (i) the 1970–76 cohorts and (ii) the 1977–1983 cohorts only. See Appendices A.1 and A.4.

---

[8]To the extent that children start school at later ages, we are underestimating the impact of the reform, as some treated individuals might be classified as control and *vice versa*.

Following Chicoine (2019), the equations in System 1 summarise the reform intensity variable in district d_n for gender s_i; $n \in \{1, ..., 103\}$. Reform intensity can be interpreted as the maximum number of additional years of schooling for each gender in district d that can ensue from the elimination of primary school fees. Thus, the theoretical magnitude ranges from 0 to 7 and is inversely related to pre-reform educational performance.

$$Intensity_{c,d_n,s_i} = \begin{cases} \sum_{g=0}^{7}(7-g) \times F_{g,d,s} & \text{if } c \geq 1995 \\ \sum_{g=1995-c}^{7}(7-g) \times F_{g,d,s} & \text{if } 1989 \leq c \leq 1994 \\ 0 & \text{if } c \leq 1988 \end{cases} \quad (1)$$

where c is for year of birth and g is for years of primary education (i.e., 0 to 7). F_{g,d_n,s_i} is defined as in Chicoine (2019). In each district d_n, some pre-reform gendered fraction F of individuals has never attended school: F_{0,d_n,s_i}. Intuitively, these individuals would have benefited the most from the elimination of fees, as they had a seven-year gap to remedy, wherefrom the pre-multiplication of F by $7 - g = 7$. Moreover, some fraction of individuals of gender s_i in district d_n have only completed 1 year of primary education: F_{1,d_n,s_i}. Analogously, F_{7,d_n,s_i} denotes the fraction of people of gender s_i who have completed primary education in district d_n. This fraction of the population would have had nothing to gain from the removal of fees, and this is reflected by the $7 - g$ pre-multiplication, whereby g is 7 in this case. Figure 3 graphs the national average of the reform intensity variable per gender and year-of-birth cohort for the analysis sample.

Figure 3: Reform Intensity by Gender and Birth Cohort
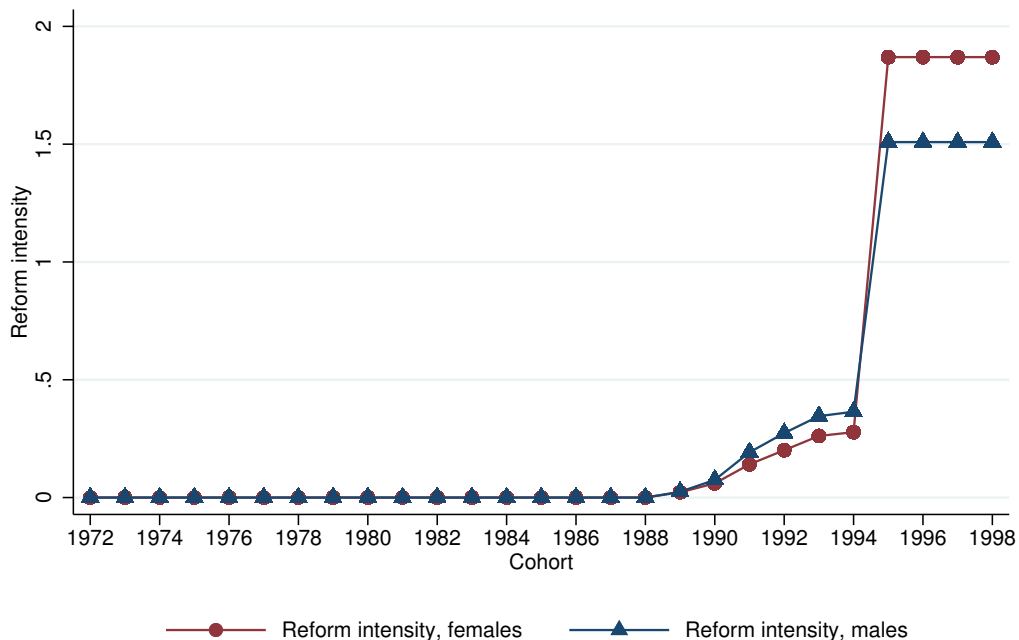
Census 2002 data is used to gauge the potential impact (reform intensity) of free primary education. Reform intensity is computed based on the educational performance of individuals aged 19–32 at the time of the reform (14 birth-year cohorts, 1970–83). Reform intensity is district-gender-birth-cohort specific. We take means over districts per gender and birth cohort. The analysis sample is comprised of the graphed cohorts. 1989 is the first partially treated cohort and 1995 is the first fully treated cohort.

4.3 Model Specification and Assumptions

The design of the intensity variable leads to a difference-in-differences framework whereby age groups are the units of observation and the year-of-birth cohort, gender and the district of residence inform the intensity of treatment.[9] All age groups are control units in 2002. In 2012, some remain control (ages 24 to 30, born 1982–88), others become partially (ages 18 to 23, born 1989-94) or fully treated (ages 14 to 17, born 1995–98).[10]

Moreover, because the analysis relies on age groups, year-of-birth cohorts and period information, the literature that deals specifically with age-period-cohort (APC) analyses further informs the setup of our empirical model. APC analyses have to overcome a specific identification issue that is due to the fact that any one variable among the age, cohort and period effects can be determined as a linear combination of the remaining two. The age, period and cohort controls are all meant to act as proxies for variables that are relevant to the empirical model but which are not observable. These underlying variables are not themselves linearly dependent (Heckman & Robb, 1985).

Heckman & Robb (1985) mention that the simplest solution is to assume that one of the age, period or cohort effects is zero. This strategy comes at an important cost if the assumption is wrong. However, due to the peculiarities of this study, we argue that this risk is acceptable. Unlike APC studies, we are not concerned with the exact magnitude of the age, cohort or period effects. We do not interpret these effects or base any conclusions on their magnitude. Instead, our interest lies with the impact of the primary school reform. Therefore, we run specifications that confront all possible specifications: *(1) cohort effects are zero, (2) age effects are zero, and (3) year effects are zero*. Namely, Equations 2, 3 and 4, respectively, to which district and gender effects are also added.

$$Education_{a,p,d,s} = \beta_1 Intensity_{a,p,d,s} + \gamma_{1,a} + \tau_{1,p} + \delta_{1,d} + \eta_{1,s} + \epsilon_{1,a,p,d,s} \tag{2}$$

$$Education_{c,p,d,s} = \beta_2 Intensity_{c,d,s} + \rho_{2,c} + \tau_{2,p} + \delta_{2,d} + \eta_{2,s} + \epsilon_{2,c,p,d,s} \tag{3}$$

$$Education_{a,c,d,s} = \beta_3 Intensity_{c,d,s} + \gamma_{3,a} + \rho_{3,c} + \delta_{3,d} + \eta_{3,s} + \epsilon_{3,a,c,d,s} \tag{4}$$

where $Intensity$ is the interaction variable of a typical difference-in-differences framework. a stands for age (14 to 30 year olds), c is for cohort (1972 to 1998), p is for period (2002, 2012), d denotes the district of observation (103 districts), and s is for gender. $Education$ can be any outcome of interest: enrolment rates or average years of primary education.

We expand the above specifications with additional level effects in the form of interactions. Since the age and period variables define the panel, we focus on these and assume cohort effects are not significant. Then, the strategy is to include district-period, age-period and age-district effects progressively for specifications *(4), (5)* and *(6)*, as summarised in Equation 5. The objective is to account for trends and further heterogeneities in the data.

---

[9]The census dataset does not include a variable documenting the district of birth of individuals. We conduct a sensitivity analysis in this regard. See Appendix A.5.

[10]See Appendix A.3 for an age-cohort-period table.

$$Education_{a,p,d,s}=\beta_4 Intensity_{a,p,d,s}+\gamma_{4,a}+\tau_{4,p}+\delta_{4,d}+\eta_{4,s}+\pi_{4,a,p}+\omega_{4,a,d}+\theta_{4,p,d}+\epsilon_{4,a,p,d,s} \quad (5)$$

The resulting framework is a high-dimensional fixed effects model. As the treatment intensity variable is computed at district level, then clustering at the district level is advised (Abadie et al., 2017). Even though treatment is also gender-specific, gender values will be correlated within districts. Moreover, higher levels of aggregation are also preferred for clustering because they are more conservative. Lucas & Mbiti (2012b,a) and Chicoine (2019) use the same level of clustering.

The identification assumption per which cohort effects are zero, as in Equation 5, has been commonly employed in the literature. In the case of Krueger & Pischke (1992), the authors argue that since their cohorts of interest are close together, they must be similar, and thus cohort effects are assumed zero. Their study is of a pension reform and its impacts on labour force participation. Similarly, Machin et al. (2011) also explore an age-year panel in their study of the impact of compulsory school reforms on crime. They control for age and year effects, and assume cohort effects are zero. In addition, we assume all relevant time-varying covariates are captured by the $Period$ dummy and the interaction terms $Age \times Period$ and $District \times Period$. Unobserved time-fixed variables are eliminated by the employment of the various layers of fixed effects that capture age, district, age-district and gender-specific characteristics.

The treatment effect is described in Equation 6. The second term of this equation, i.e., the difference between the pre- and post-reform outcomes of the control group had it been treated, is the counterfactual that is not observed.

$$\hat{\beta} = E\big(Education_{a,2012,d,s}^{Treated} - Education_{a,2002,d,s}^{Treated} \mid Intensity_{a,2012,d,s} > 0, \gamma, \delta, \eta, \pi, \omega, \theta\big)-$$

$$-E\big(Education_{a,2012,d,s}^{Control} - Education_{a,2002,d,s}^{Control} \mid Intensity_{a,2012,d,s} > 0, \gamma, \delta, \eta, \pi, \omega, \theta\big) \quad (6)$$

The difference-in-differences estimation method assumes that the counterfactual is equal to the difference between the pre- and post-reform outcomes of the control group in the absence of treatment. This is the parallel trends assumption, presented in Equation 7.

$$E\big(Education_{a,2012,d,s}^{Control} - Education_{a,2002,d,s}^{Control} \mid Intensity_{a,2012,d,s} > 0, \gamma, \delta, \eta, \pi, \omega, \theta\big) =$$

$$= E\big(Education_{a,2012,d,s}^{Control} - Education_{a,2002,d,s}^{Control} \mid Intensity_{a,2012,d,s} = 0, \gamma, \delta, \eta, \pi, \omega, \theta\big) \quad (7)$$

If all identification assumptions are satisfied, then the difference-in-differences estimation method identifies the *average treatment on the treated effect* presented in Equation 8.

$$\hat{\beta}_{DiD} = E\big(Education_{a,2012,d,s}^{Treated} - Education_{a,2002,d,s}^{Treated} \mid Intensity_{a,2012,d,s} > 0, \gamma, \delta, \eta, \pi, \omega, \theta\big)-$$

$$-E\big(Education_{a,2012,d,s}^{Control} - Education_{a,2002,d,s}^{Control} \mid Intensity_{a,2012,d,s} = 0, \gamma, \delta, \eta, \pi, \omega, \theta\big) \quad (8)$$

Figure 4 brings strong evidence in support of the parallel trends assumption. First, it is visible that among the control units, there is no significant difference between the out-of-school gendered ratios of 2002 and 2012, i.e., the dotted and solid lines either overlap or are parallel for each gender if the line marker is c. Second, starting with ages 23–24, when cohorts are classified as partially treated in 2012, it becomes apparent that the 2012 outcomes diverge from their 2002 counterparts by following a downward path.

Figure 4: Parallel Trends Assumption, Never-Enrolled Population Ratios
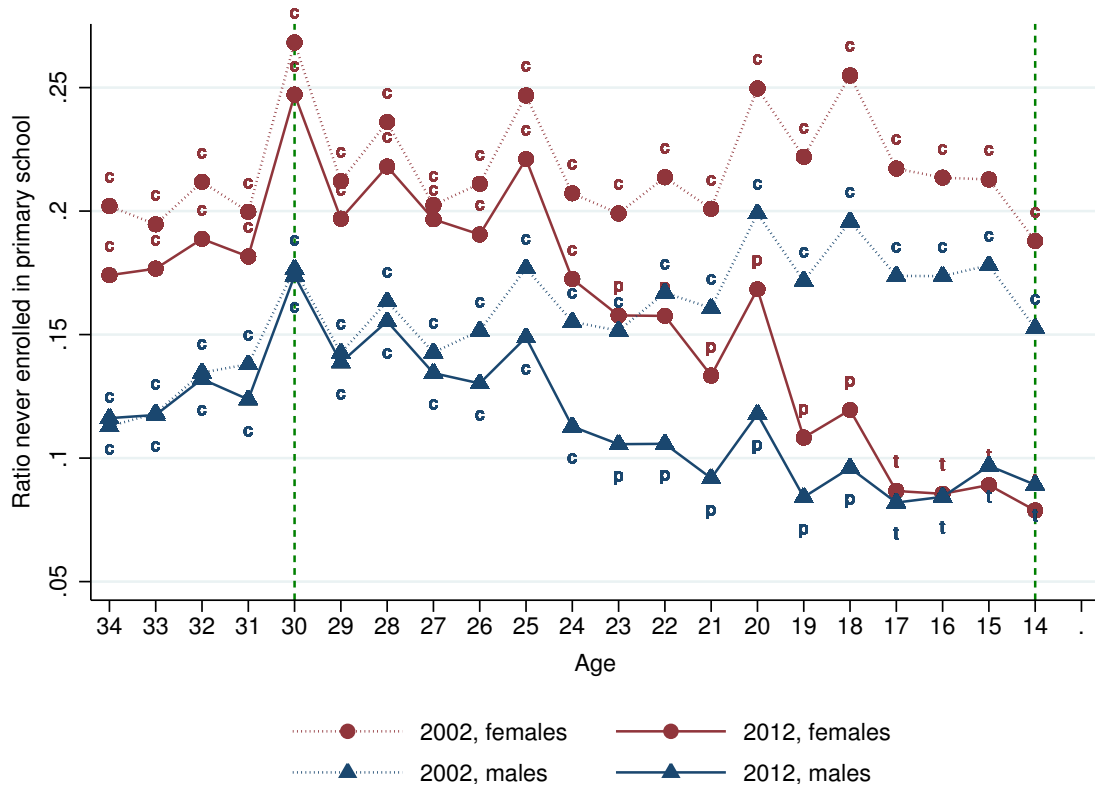
The vertical, dotted, green lines show the first and last age cohorts to be included in the analysis sample, 14 and 30 years of age, respectively. c is control. p and t stand for partial and full treatment, respectively.

Moreover, Figure 4 brings suggestive evidence of the impact of school fee elimination on the percentage of individuals who have never attended school; solid lines are notably below the dotted lines for each gender starting with the partially treated cohorts. Additionally, the graph also shows the gender imbalance regarding educational outcomes. A higher share of control women have never attended school compared to males; red lines are substantially above the blue lines for control cohorts. The graph suggests, however, that the difference between treated females and males has been eliminated following the reform; solid lines converge for the treated age cohorts.

4.4 Impact Heterogeneities

To assess the heterogeneity of reform impacts regarding gender and colonial infrastructure, we interact the reform variable with a variable capturing gender and the standardised number of schools per 100,000 children aged 7–13 in 2012. See Equations 9 and 10.

16

$$Education_{a,p,d,s} = \beta_5 Intensity_{a,p,d,s} + \alpha_5 Intensity_{a,p,d,s} \times Female+$$

$$\gamma_{5,a} + \tau_{5,p} + \delta_{5,d} + \eta_{5,s} + \pi_{5,a,p} + \omega_{5,a,d} + \theta_{5,p,d} + \epsilon_{5,a,p,d,s} \qquad (9)$$

$$Education_{a,p,d,s} = \beta_6 Intensity_{a,p,d,s} + \alpha_6 Intensity_{a,p,d,s} \times Female+$$

$$+\sum_{i=1}^{3} \phi_i Intensity_{a,p,d,s} \times Schools_{i,d} + \sum_{i=1}^{3} \zeta_i Intensity_{a,p,d,s} \times Schools_{i,d} \times Female+$$

$$+\gamma_{6,a} + \tau_{6,p} + \delta_{6,d} + \eta_{6,s} + \pi_{6,a,p} + \omega_{6,a,d} + \theta_{6,p,d} + \epsilon_{6,a,p,d,s} \qquad (10)$$

Where $i \in \{1, 2, 3\}$ stands for Catholic, Protestant and village-authority schools.

We do not suspect the variable measuring colonial infrastructure to be endogenous because the fixed effects that we employ take care of the factors that explain the location of colonial schools, which might also be correlated with present-day outcomes. These variables can be altitude, the weather and climate, disease prevalence, access to water and roads at the time of the missionary influx. (Nunn, 2014a; Huillery, 2009; Montgomery, 2017).

5 Results

5.1 Impact of School Fee Elimination

Among the fully treated cohorts, reform intensity is 1.7 years on average and the standard deviation is 0.9 years. The minimum value is 0.2 and the maximum is 5.1 years. We find that the implementation of the reform has improved average educational outcomes. See Table 4. A one-year increase in reform intensity triggers a reduction of 4 percentage points in the ratio of individuals who never enrol in education and a 0.29 increase in average years of primary education, i.e., 3.5 months. At the average intensity of 1.7 years of free education, the effect is a decrease of 6.8 percentage points in the never-enrolled population and an increase of 6 months in average primary school achievement. These are the results put forward for each block of dependent variables in column (6), which is the most comprehensive and our preferred specification. The other specifications also point to significant coefficients: a 4–5 percentage-point decrease in the ratio of individuals who have never enrolled in school, and a 0.26–0.44 years increase in average primary education for one additional year of free education.

We note, however, that if educational outcomes are only averaged over populations who *have* enrolled in primary education—third block of variables in Table 4—, then the afore-mentioned positive effects are no longer strong or consistent. These results suggest that the reform has mainly improved enrolment rates, which in return have boosted average educational levels. As the percentage of people who have never enrolled in primary education is reduced, the average education of the concerned population will automatically increase, and the magnitude of this increase would have been amplified by the efficacy of the reform in improving the educational outcomes of those enrolled.

Table 4: Impacts of School Fee Elimination on Educational Outcomes

Explanatory variable	RATIO OF NEVER-ENROLLED INDIVIDUALS - All individuals -						YEARS OF PRIMARY EDUCATION - All individuals -						YEARS OF PRIMARY EDUCATION - Only enrollees -					
	(1)	(2)	(3)	(4)	(5)	(6)	(1)	(2)	(3)	(4)	(5)	(6)	(1)	(2)	(3)	(4)	(5)	(6)
Reform intensity	-0.05***	-0.05***	-0.05***	-0.04***	-0.04***	-0.04***	0.44***	0.34***	0.34***	0.41***	0.26***	0.29***	0.17***	-0.01	-0.01	0.18***	-0.03***	0.07***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Fixed Effects																		
District	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Gender	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth cohort	-	Yes	Yes	-	-	-	-	Yes	Yes	-	-	-	-	Yes	Yes	-	-	-
Age	Yes	-	Yes	Yes	Yes	Yes	Yes	-	Yes	Yes	Yes	Yes	Yes	-	Yes	Yes	Yes	Yes
Period	Yes	Yes	-	Yes	Yes	Yes	Yes	Yes	-	Yes	Yes	Yes	Yes	Yes	-	Yes	Yes	Yes
District × Period	-	-	-	Yes	Yes	Yes	-	-	-	Yes	Yes	Yes	-	-	-	Yes	Yes	Yes
Age × Period	-	-	-	-	Yes	Yes	-	-	-	-	Yes	Yes	-	-	-	-	Yes	Yes
Age × District	-	-	-	-	-	Yes	-	-	-	-	-	Yes	-	-	-	-	-	Yes
Within R^2	0.19	0.13	0.13	0.17	0.09	0.07	0.26	0.08	0.09	0.26	0.06	0.05	0.17	0.00	0.00	0.18	0.00	0.02
F	516	263	263	548	294	199	621	165	165	615	169	174	151	2	2	154	12	22
Nr. clusters	103	103	103	103	103	103	103	103	103	103	103	103	103	103	103	103	103	103
N	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Errors are clustered at the level of districts and are presented in parenthesis. All regressions include the 14–30 age groups of 2002 and 2012. In terms of birth cohorts, this means: 1972–98. The 1989–94 cohorts have been partially treated, and the 1995–97 cohorts have been fully treated. Census 2002 data is used to gauge reform intensity. Reform intensity is computed based on the educational performance of individuals who were aged 19–32 at the time of the reform (14 birth-year cohorts, 1970–83). Reform intensity is district-gender-birth-cohort specific. The sample only includes the districts of Mainland Tanzania as of 1988.

We tentatively conclude that fees have been an important obstacle in the achievement of full enrolment; however, fees do not appear to have been the main problem behind dropouts for those who do enrol in primary education. An alternative explanation is that the average educational achievement of those who would have enrolled in school regardless of the reform may have been watered down by the performance of individuals who enrol as a consequence of the removal of fees but for whom the elimination of fees is insufficient to allow them to complete primary education, and thus they drop out. Consequently, the latter group of students cancels any improvement in grade achievement for those who would have enrolled anyway, fees or no fees.

5.2 Impact Heterogeneity, Gender, Colonial Infrastructure

In Table 5, we assess the heterogeneity of the reform by estimating Equations 9 and 10. There is strong evidence that reform impacts are heterogeneous in terms of gender and the colonial infrastructure of the 1940s. The impact for females was at least double the effect for males. Specification (1) in Table 5 shows that while the impact of the reform on the never-enrolled ratio of males was 2 percentage points on average, the magnitude for females stood at 4 percentage points for one additional year of free primary education. As pointed out in Section 3, this may reflect the fact that there are more females who have never attended school than there are males. Thus, there is more room for females to improve relative to male cohorts. Similarly, the effect is also carried over when average educational achievement is the dependent variable. For males, the reform triggered a 0.10 increase for one additional year of free education. For females, the effect was 0.17 of a year higher. Finally, Table 5 also suggests that the reform has improved the educational achievement of enrolled females. Although the effect is small, it is statistically significant. We test if the sum of the coefficients in column (1) is statistically different from zero. The F-test rejects the null of non-significance with a p-value of 0. In contrast, the result for males is not robust. Overall, the results of specification (1) support the hypothesis that the elimination of fees has reduced the educational gender gap.

Table 5 also brings evidence that the reform has further reduced the educational gender gap in favour of females residing in districts that have benefited from stronger investments in missionary or local-authority schools during colonial rule. In contrast, these historical institutions do not appear to suggest any reform heterogeneity for males. For a standard deviation increase in the Catholic school infrastructure, the elimination of school fees further improves the enrolment of females by 1 percentage point. The same applies to the Protestant infrastructure. Similarly, the impact of fee removal on the average number of years of primary education is increased by 0.05 and 0.03 for one standard deviation increase in the scaled number of Catholic and Protestant schools, respectively. Specification (5) shows that for a deviation increase in the number of both Protestant and Catholic schools per 100,000 children, the baseline reform impacts for females, i.e., 5 percentage points higher enrolments and 0.28 better education for one additional year of free education, are further increased to 7 percentage points and 0.37 years of education.

Table 5: Heterogeneous Reform Impacts, Gender and Colonial Infrastructure

Explanatory variables	RATIO OF NEVER-ENROLLED INDIVIDUALS - All individuals -					YEARS OF PRIMARY EDUCATION - All individuals -					YEARS OF PRIMARY EDUCATION - Only enrollees -				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Reform intensity (years)	-0.02***	-0.02***	-0.03***	-0.02***	-0.03***	0.10**	0.11***	0.10***	0.10***	0.11***	0.02	0.03	0.02	0.02	0.03
	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Reform × Female	-0.02***	-0.02***	-0.02***	-0.02***	-0.02***	0.17***	0.18***	0.18***	0.18***	0.18***	0.04***	0.04***	0.04***	0.04***	0.04***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Reform × Catholic		0.00			0.00		0.00			0.01		0.02			0.02
		(0.00)			(0.00)		(0.02)			(0.02)		(0.01)			(0.01)
Reform × Female× Catholic		-0.01***			-0.01***		0.05***			0.05***		-0.00			-0.00
		(0.00)			(0.00)		(0.02)			(0.02)		(0.00)			(0.00)
Reform × Protestant			0.00		0.00			-0.01		-0.01			0.00		-0.00
			(0.00)		(0.00)			(0.03)		(0.03)			(0.01)		(0.01)
Reform × Female × Protestant			-0.01***		-0.01***			0.04***		0.03**			0.00		0.00
			(0.00)		(0.00)			(0.01)		(0.01)			(0.00)		(0.00)
Reform × Village authority				-0.00	-0.00				0.01	0.01				-0.01	-0.01
				(0.00)	(0.00)				(0.01)	(0.01)				(0.01)	(0.01)
Reform × Female × Village				-0.00	-0.00				0.02	0.01				0.01***	0.01***
				(0.00)	(0.00)				(0.02)	(0.01)				(0.00)	(0.00)
Adjusted within R^2	0.09	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.03	0.03	0.03	0.03	0.03
F	258	134	118	147	69	233	136	113	132	68	59	36	31	33	19
Nr. clusters	103	103	103	103	103	103	103	103	103	103	103	103	103	103	103
N	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Errors are clustered at the level of districts and are presented in parenthesis. All regressions are based on the preferred specification, which includes the following fixed effects: District, Gender, Age, Period, District × Period, Age × Period, Age × District. The Catholic, Protestant and village-authority variables are standardised values of the normalised number of schools per 100,000 children aged 7–13 in 2012; $\mu = 1$ and $\sigma = 1$. All regressions include the 14–30 age groups of 2002 and 2012. In terms of birth cohorts, this means: 1972–98. The 1989–94 cohorts have been partially treated, and the 1995–97 cohorts have been fully treated. Census 2002 data is used to gauge reform intensity. Reform intensity is computed based on the educational performance of individuals who were aged 19–32 at the time of the reform (born 1970–83). The sample includes the districts of Mainland Tanzania as of 1988.

If average education is computed among enrollees only, then there is no heterogeneity in terms of the missionary infrastructure, but there is some heterogeneity regarding village-authority schools, as they favour larger reform impacts for females.

These findings agree with the historical facts presented in Section 2, per which the missionary infrastructure was mainly interested in increasing the number of religious converts and less concerned about the quality of the education they offered. Consequently, their infrastructure was larger relative to government schools, and their teaching less rigorous. It appears that this legacy has persisted to this day, as districts with historically larger missionary investments perform better in terms of the female-male enrolment gaps—presumably because they have more schools and thus are better able to relax the time constraints that limit the educational opportunities of girls.

Table 6: Educational Gender Gap and Colonial Infrastrucutre

	RATIO NEVER ENROLLED	AVERAGE EDUCATION - All individuals -	AVERAGE EDUCATION - Only enrollees -
	(1)	(2)	(3)
Female	0.05***	-0.23***	0.09***
	(0.00)	(0.02)	(0.01)
Female × Catholic	-0.01***	0.04**	-0.00
	(0.00)	(0.01)	(0.00)
Female × Protestant	-0.01***	0.07***	0.01**
	(0.00)	(0.02)	(0.00)
Female × Village authority	-0.00	0.02	0.01***
	(0.00)	(0.02)	(0.00)
Adjusted within R^2	0.35	0.20	0.18
F	46	24	138
Nr. clusters	103	103	103
N	7,004	7,004	7,004

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Errors are clustered at the level of districts. All regressions include the following fixed effects: District, Gender, Age, Period, District × Period, Age × Period, Age × District. The Catholic, Protestant and village authority variables are standardised values of the normalised number of schools per 100,000 children aged 7–13 in 2012; $\mu = 1$ and $\sigma = 1$.

Columns (1) and (2) of Table 6 show that, on average, there is a 5 percent gender gap in enrolment and a 0.23 gap in primary educational achievement, i.e., 3 months, with females lagging behind. However, once enrolled, females accumulate more education. The interaction terms bring additional information. We learn that the gap is smaller in districts where early investments in education were larger. For instance, the enrolment gap is reduced to 3 percent in districts endowed with an additional standard deviation in the density of Catholic and Protestant schools. Similarly, instead of a 3-month gap, in these districts, the gap would be of 1.4 months. Furthermore, column (3) of Table 6 shows that among the enrolled, females stay in school one month longer than males on average. Their advantage is further increased in areas with colonial-time Protestant and government investments. This heterogeneity is similar to that presented in Table

21

5. In addition, column (3) also vindicates the findings of Nunn (2014a) and Becker & Woessmann (2008), who argue that Protestant missions were more preoccupied to promote literate girls relative to Catholic schools.

Tables 5 and 6 suggest that gender gaps were already smaller in districts with larger colonial infrastructure, and the reform has also been more effective at reducing the educational gender gap in these districts. The reform continues to add to the advantages created by early investments in education and perpetuates historical legacies. All groups have received a boost, especially females. However, the females that were already better off have benefited more than the females impacted by poor historical investments in education.

5.3   Mechanism

Inspired by the work of Huillery (2009), we investigate whether the persistence of investments in infrastructure can explain why current reforms perpetuate the legacies of the past. We suspect that districts where investments in education took place relatively early continue to invest more than other districts.

Table 7: Educational Gender Gap and Current Infrastrucutre

|	RATIO NEVER ENROLLED	AVERAGE EDUCATION - All individuals -	AVERAGE EDUCATION - Only enrollees -
	(1)	(2)	(3)
Female	0.05***	-0.23***	0.09***
	(0.00)	(0.02)	(0.01)
Female $\times$ Current infrastructure	-0.01***	0.08***	0.01*
	(0.00)	(0.02)	(0.01)
Adjusted within R^2	0.35	0.20	0.18
F	89	48	124
Nr. clusters	103	103	103
N	7,004	7,004	7,004

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Errors are clustered at the level of districts. All regressions includes the following fixed effects: District, Gender, Age, Period, District $\times$ Period, Age $\times$ Period, Age $\times$ District. Current infrastructure consists of the 2012 number of school per district scaled to the 2012 population of children aged 7–13 and standardised such that $\mu = 1$ and $\sigma = 1$.

Against the background of the initiative *"Big Results Now in Education"*, the government has committed to sharing information about the performance of schools to promote accountability (Cilliers et al., 2020). We use the output of this initiative, i.e., a publicly available list of the universe of schools, to investigate the persistence of investments in schools. This list includes the GPS coordinates of schools and performance indicators.

Table 7 shows that a larger school infrastructure nowadays is also associated with a smaller gender gap. A more readily available infrastructure is arguably capable of relaxing binding household constraints, which generally keep girls away from school. Distance

to school is a serious deterrent of school attendance in developing countries in general (Muralidharan & Prakash, 2017) and in Tanzania, especially for girls and for children in rural areas (Al-Samarrai & Reilly, 2000; Kondylis & Manacorda, 2012). Moreover, the same objective constraint regarding school availability may matter more for Tanzanian girls than for boys (Lihwa et al., 2019), as the demands for female and male time are different and thus, so are their opportunity costs. Mason & Khandker (1996) argue that the time opportunity cost for girls of primary school age in Tanzania is larger than for boys. Consequently, the remoteness of schools is more costly for girls than it is for boys.

Figure 5: Persistence of Investments in Education

1947 Catholic, Protestant and village schools per 100,000 children aged 7–13 in 2012
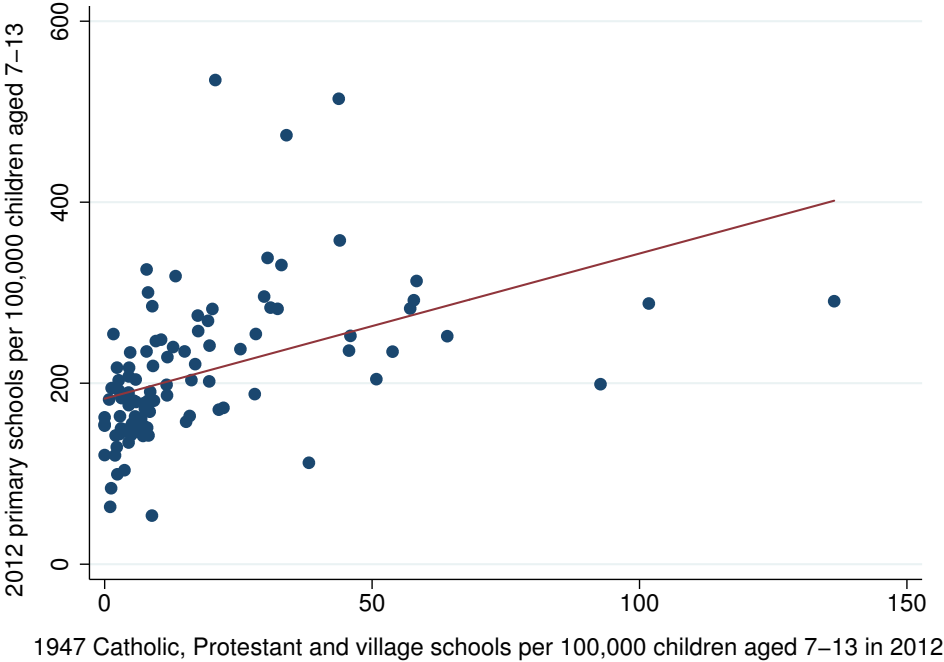
The 1947 schools are added together, i.e., Catholic, Protestant and village-authority schools. Both axes are scaled to the population of children aged 7–13 in 2012.

Figure 5 brings suggestive evidence that past and current infrastructure are related. The 1947 school infrastructure per 100,000 students aged 7–13 is highly correlated with the infrastructure in 2012. The correlation coefficient is 0.45. This evidence does not reflect a causal relationship, and the reader should treat the information accordingly. Table 7 and Figure 5 bring suggestive evidence that the mechanism which has enabled the elimination of school fees to disproportionately reach girls in districts with a stronger history of educational investments is the inherited proclivity of these districts to invest more in their educational infrastructure. Consequently, districts with above-average school densities are readier to absorb the increased demand for education in the aftermath of reforms. If schools are not sufficiently available, then the impacts of school fee elimination may be dampened by binding constraints such as limited infrastructure.

6 Robustness Checks

6.1 Falsification Test: Timing of the Reform

To scrutinise the results of Table 4, we falsify the treatment variable by intentionally misplacing the timing of the reform. This is the same approach adopted by Chicoine (2019). The true timing is 2002, however, we run iterations whereby the reform is assumed to have been implemented in each of the years between 1992 and 2005.[11]

Figure 6: Falsification Analysis
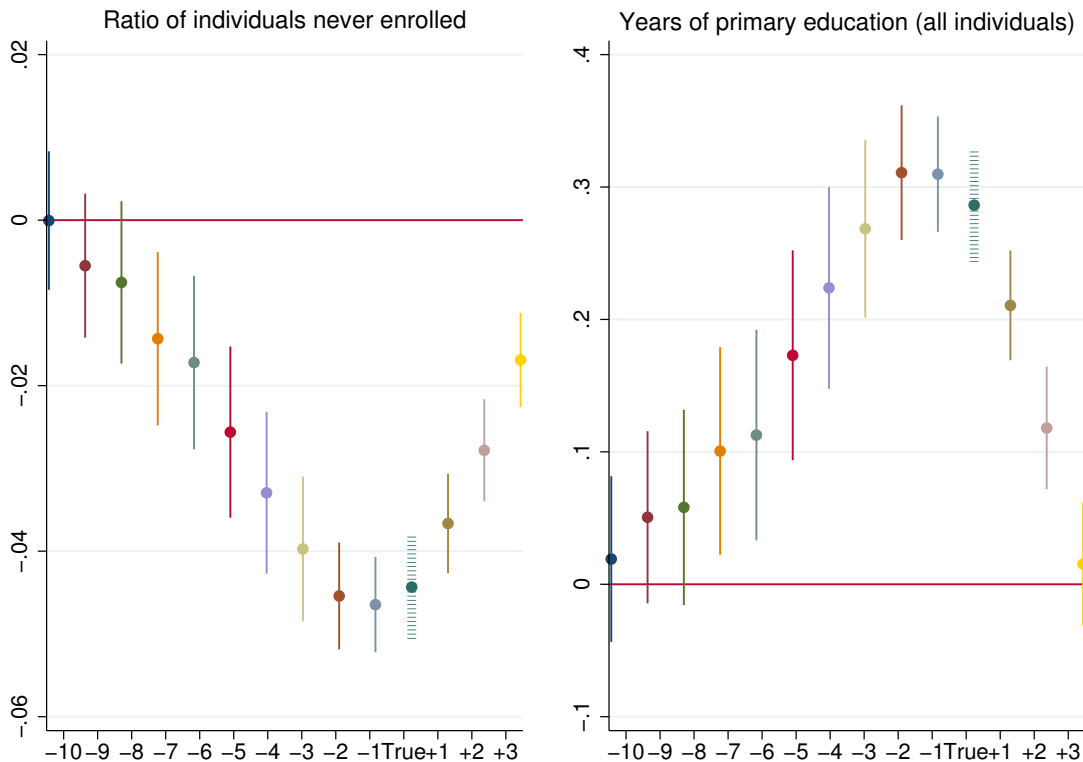
Point estimates are accompanied by 95% confidence intervals. The reform impact coefficient is plotted for 14 regressions. Control and treatment cohorts are pushed forward or backward each time by 1 year.

The results of the falsified regressions are plotted in Figure 6. All regressions are based on the preferred specification, which includes the following fixed effects: District, Gender, Age, Period, District × Period, Age × Period, Age × District. This figure shows how both the magnitude and significance of the estimated impact are strengthened the closer the falsified timing of the reform is to 2002, the true implementation year. The fact that the magnitude is slightly bigger in $T - 1$ may be due to the fact that the elimination of fees was first announced in 2001 and then implemented in 2002. Arguably, this has allowed parents to send their children to school earlier, in an attempt to avoid enrolment refusals once educational facilities became overcrowded in the aftermath of the reform. Moreover, late enrolment is a chronic problem in Tanzania (Mason & Khandker, 1996). We may be slightly underestimating the impact of the reform because we assume that all children enrol at seven years of age. If a large number of students start school at age

---

[11]Appendix A.3 shows which cohorts are considered for each falsification test.

eight, then this can show up in the data as a stronger reform impact at $T - 1$. Overall, we are confident that the falsification test is supportive of our analysis.

6.2   Reform Measurement

The benchmark reform intensity is gauged based on the educational performance of all individuals born between 1970 and 1983, aged 19–32 at the time of the reform in 2002.[12] To explore the robustness of our findings, we use two different reform variables. Instead of focusing on all 14 cohorts between 1970 and 1983, we take the 7 most recent cohorts, 1977–83, and then the 7 oldest cohorts, 1970–76.[13] Appendix A.4 shows that the benchmark results are robust to these alternative specifications of reform intensity. The magnitude of coefficients and the qualitative implications of Table 4 are maintained.

6.3   Analysis Sample

In the benchmark analysis, we have relied on an individual's district of residence, as opposed to district of birth, to compute the district-aggregated educational outcomes. Therefore, the dependent variable is likely to suffer from measurement error. Importantly, this measurement error may be non-random if there is a tendency for some districts to attract individuals that are better-educated. Unfortunately, we do not have data on district of birth. To mitigate some of the measurement error, we exclude from the computation of the aggregated outcomes of interest those individuals for whom the region of residence does not match their region of birth. Admittedly, this solution is inferior to the one whereby district mismatches are removed as district data is more granular.[14] Nevertheless, districts within regions are similar and share commonalities. Thus, we argue this correction meets its intended purpose. Consequently, we remove 18 percent of the individual-level sample, as these individuals have moved regions since birth. The aggregated variables are then re-computed. The new regression results support our previous findings. See Appendix A.5. Differences in magnitude between Tables 5 and Appendix A.5 are negligible, although we do gain in precision for some of the estimates if the variable of interest is the education of enrollees. The impact of the reform on males is occasionally significant and more so in districts with historically larger Catholic missionary presence.

---

[12]We start at age 19 because we want to avoid any issues created by late enrolments and their ensuing late graduation. This could be a problem because the scope of the reform intensity variable is to describe pre-reform performance. If cohorts include individuals who might have been impacted by the reform, then this would affect our identification strategy. We do not go beyond the 1970 cohort because, against the background of rapid and important changes in Tanzania post-independence, older cohorts are too detached from Tanzania's educational situation after 1980.

[13]Appendix A.1 compares these three cohort-based measurements, 1970–83, 1970–76 and 1977–83. The intuition is that pre-reform educational achievement and reform intensity are inversely related. While the educational situation of females has not changed significantly across the 1970–83 cohorts, for males, however, the reform intensity variable suggests a marked improvement. For instance, the variable is lower if the 1977–83 cohorts are considered, which is a consequence of these cohorts' superior educational results relative to their older counterparts. Similarly, the variable reaches its highest magnitude if the 1970–76 cohorts are employed, as they had more room for improvement, and reform intensity is thus estimated to be stronger. Benchmark intensity sits midway between the aforementioned specifications.

[14]There are 103 districts in the sample but only 18 regions.

6.4 Falsification Test: Railway Infrastructure

Finally, we check the robustness of the results presented in Section 5.2 by falsifying the type of colonial infrastructure that we consider. Instead of the educational infrastructure of 1947, we explore the 1950 railway infrastructure of Tanganyika. This type of falsification test has also been employed by Nunn (2014a). For this purpose, we have geo-referenced a map compiled by the Department of Lands and Surveys of Tanganyika around the same time when the school infrastructure map was also compiled.[15] See Appendix A.6.

Table 8: Falsification Analysis: Railway Infrastructure

	RATIO NEVER ENROLLED		AVERAGE EDUCATION - All individuals -		AVERAGE EDUCATION - Only enrollees -	
	(1)	(2)	(1)	(2)	(1)	(2)
Reform intensity (years)	-0.02***	-0.03***	0.10**	0.12***	0.02	0.03
	(0.01)	(0.01)	(0.05)	(0.04)	(0.02)	(0.02)
Reform × Female	-0.02***	-0.02***	0.17***	0.16***	0.04***	0.04***
	(0.00)	(0.00)	(0.03)	(0.02)	(0.01)	(0.01)
Reform × Railway goes through district	-0.00		-0.00		-0.01	
	(0.01)		(0.04)		(0.02)	
Reform × Female × Railway goes through district	-0.00		0.01		0.01	
	(0.01)		(0.04)		(0.01)	
Reform × Railway is close		0.01		-0.05		-0.02
		(0.01)		(0.04)		(0.02)
Reform × Female × Railway is close		-0.01		0.05		0.01
		(0.01)		(0.04)		(0.01)
Adjusted within R^2	0.09	0.09	0.10	0.10	0.03	0.03
F	137	127	121	117	30	31
Nr. clusters	103	103	103	103	103	103
N	7,004	7,004	7,004	7,004	7,004	7,004

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Errors are clustered at the level of districts and are presented in parenthesis. All regressions include the following fixed effects: District, Gender, Age, Period, District × Period, Age × Period, Age × District. The railway variables are dummies that take value 1 if the railway line intersects with the district polygon—regardless whether the intersection is notable or slight—and if the centroid of the district is within 50 km from the railway line, respectively.

Nunn (2014a) uses a dummy variable to denote whether a village or ethnic group was accessed via a railway in the early 1900s. We employ the same idea and use two dummy variables to establish whether a railway crossed the territory of any one district and whether the district centroid is within 50 kilometres from the closest railway. Table 8 shows that the school fee reform does not exhibit any heterogeneity regarding the railway infrastructure. This is reassuring and suggests that our main results are not spurious due to unobservable factors which made certain districts more attractive to colonial settlers. Our analysis thus survives the falsification test.

---

[15]The map is hosted by the Princeton University Library, https://catalog.princeton.edu/catalog/10159264.

7   Conclusion

In this study, we have compiled evidence that the elimination of primary school fees in Mainland Tanzania has improved educational outcomes on average. The fraction of individuals who have never attended school has decreased by 4 percentage points and average education has increased by 3.5 months as a result of being exposed to one additional year of free primary education. Female students, who were falling behind their male counterparts in terms of enrolment before the reform, have benefited from the elimination of school fees twice as much compared to male students. This finding is in agreement with evidence from the Ugandan fee elimination (Deininger, 2003; Grogan, 2009) but stands in contrast to the Kenyan reform (Lucas & Mbiti, 2012b). The latter reports a widening of the gender gap.

We estimate that the greatest beneficiaries of the reform have been females who reside in districts where investments in education have been relatively stronger during colonial rule. This means that females who have been disadvantaged by historically poor investments in their districts have continued to benefit less from educational reforms relative to females residing in districts with a stronger legacy of colonial schools. While we find that Protestant schools have had a stronger impact on the education of females compared to Catholic schools, we fail to reject that Catholic schools have also supported female education. Consequently, our study agrees with the findings of Nunn (2014a) and Becker & Woessmann (2008) regarding the role of Protestant schools; however, we differ from Nunn (2014a) as we argue that Catholic schools, too, have been a noteworthy vehicle in reducing gender gaps in enrolment. Lastly, our study complements the work of Montgomery (2017). The author argues that the school infrastructure of German East Africa is associated with a larger educational gender gap in present-day Tanzania. In contrast, this study brings evidence that the school infrastructure erected during the British colonial rule of Tanganyika has had the opposite effect, one whereby it has facilitated a smaller gender gap in education.

The literature has documented strong evidence that the impacts of historical events and institutions are long-lasting. We conclude that current reforms may also be perpetuating colonial legacies instead of eliminating historical disparities. Policy-makers should engage in concerted efforts to identify and address such patterns. Reforms should address historical legacies to allow the convergence of outcomes across areas with differing degrees of early investments in education. For instance, infrastructure expansions, which usually accompany the removal of school fees, should give disproportionate attention to districts that have been historically disadvantaged as a consequence of the unequal allocation of resources. Otherwise, disparities will persist, although average improvements can be registered.

A   Appendix

A.1   Reform Intensity Variations
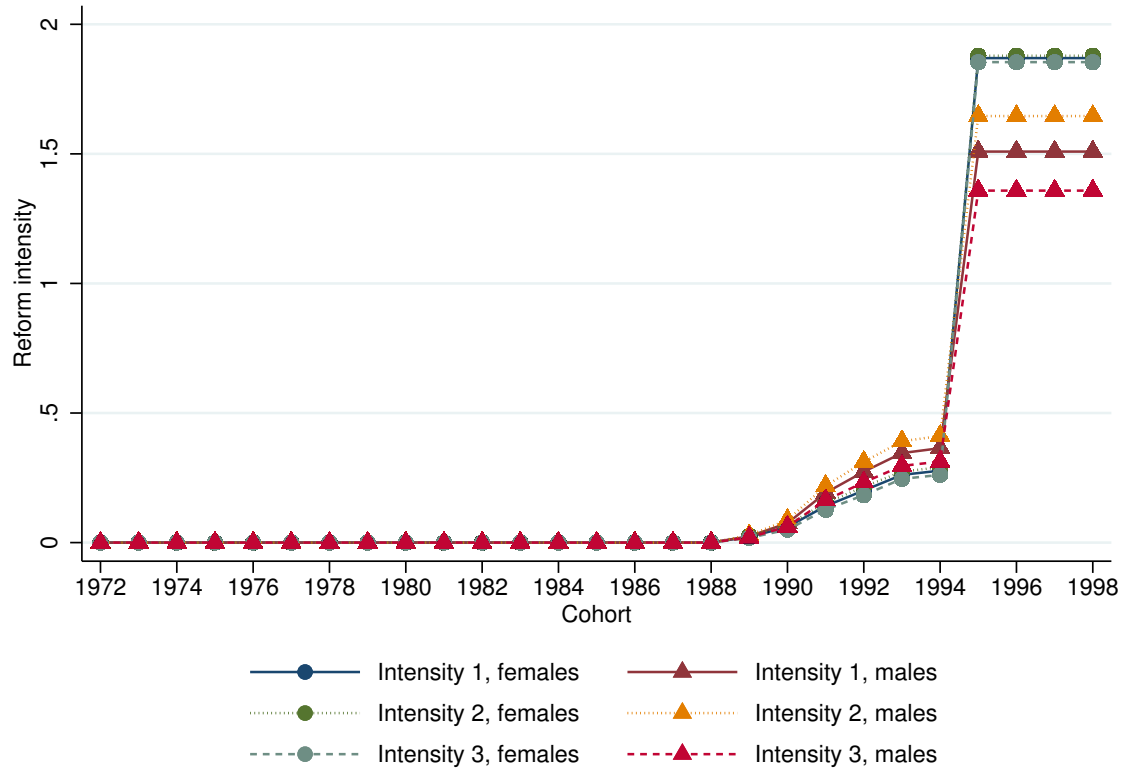
Reform intensity is computed based on the educational performance of: (**Intensity 1**) individuals who were aged 19–32 at the time of the reform (14 birth-year cohorts, 1970–83); (**Intensity 2**) individuals who were aged 19–25 at the time of the reform (7 birth-year cohorts, 1977–83); and (**Intensity 3**) individuals who were aged 26–32 at the time of the reform (7 birth-year cohorts, 1970-76). Reform intensity is district-gender-birth-cohort specific. The graph takes means over districts per gender and birth cohort. 1989 is the first partially treated cohort and 1995 is the first fully treated cohort.

A.2 District and Regional Distribution of Reform Intensity
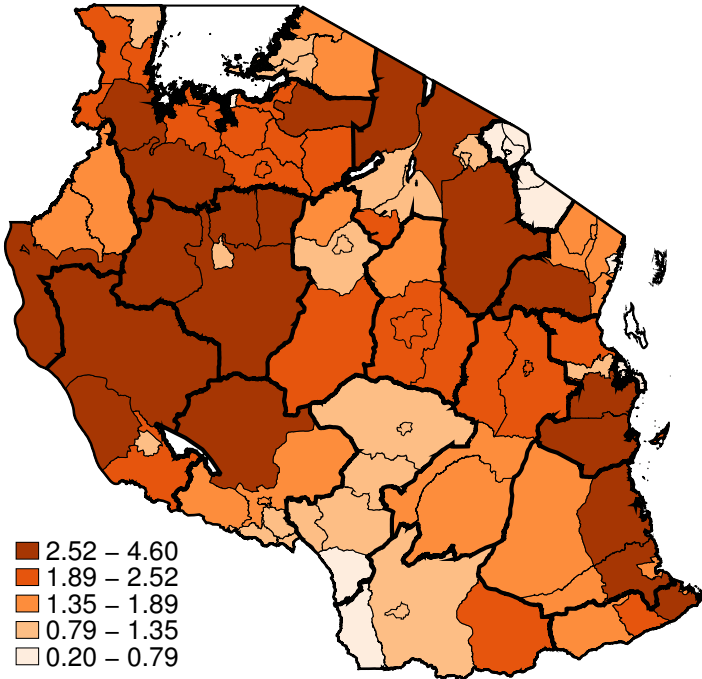
The reform intensity is computed based on the educational performance of individuals who were aged 19–32 at the time of the reform (14 birth-year cohorts, 1970–83). Reform intensity is district-gender-birth-cohort specific. Only the intensity for the fully treated cohorts is mapped. We take averages over genders. 103 Mainland districts (thin contours) and 18 regions (thick contours).

A.3 Age-Period-Cohort Table

AGE OR BIRTH COHORT

Age in 2002	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
Birth cohort (year)	91	90	89	88	87	86	85	84	83	82	81	80	79	78	77	76	75	74	73	72	71	70	69	68	67	66	65	64	63	62
True treatment status	p	p	p	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c
Age in 2012	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
Birth cohort (year)	01	00	99	98	97	96	95	94	93	92	91	90	89	88	87	86	85	84	83	82	81	80	79	78	77	76	75	74	73	72
True treatment status	t	t	t	t	t	t	t	p	p	p	p	p	p	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c
False, +3, 2002 status	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c													
False, +3, 2012 status	t	t	t	t	p	p	p	p	p	p	c	c	c	c	c	c	c													
False, +2, 2002 status		c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c												
False, +2, 2012 status		t	t	t	t	p	p	p	p	p	p	c	c	c	c	c	c	c												
False, +1, 2002 status			c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c											
False, +1, 2012 status			t	t	t	t	p	p	p	p	p	p	c	c	c	c	c	c	c											
False, -1, 2002 status					c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c									
False, -1, 2012 status					t	t	t	t	p	p	p	p	p	p	c	c	c	c	c	c	c									
False, -2, 2002 status						c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c								
False, -2, 2012 status						t	t	t	t	p	p	p	p	p	p	c	c	c	c	c	c	c								
False, -3, 2002 status							c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c							
False, -3, 2012 status							t	t	t	t	p	p	p	p	p	p	c	c	c	c	c	c	c							
False, -4, 2002 status								c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c						
False, -4, 2012 status								t	t	t	t	p	p	p	p	p	p	c	c	c	c	c	c	c						
False, -5, 2002 status									c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c					
False, -5, 2012 status									t	t	t	t	p	p	p	p	p	p	c	c	c	c	c	c	c					
False, -6, 2002 status										c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c				
False, -6, 2012 status										t	t	t	t	p	p	p	p	p	p	c	c	c	c	c	c	c				
False, -7, 2002 status											c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c			
False, -7, 2012 status											t	t	t	t	p	p	p	p	p	p	c	c	c	c	c	c	c			
False, -8, 2002 status												c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c		
False, -8, 2012 status												t	t	t	t	p	p	p	p	p	p	c	c	c	c	c	c	c		
False, -9, 2002 status													c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	
False, -9, 2012 status													t	t	t	t	p	p	p	p	p	p	c	c	c	c	c	c	c	
False, -10, 2002 status														c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c
False, -10, 2012 status														t	t	t	t	p	p	p	p	p	p	c	c	c	c	c	c	c

Where t denotes the treated, p is for partial treatment and c is for control.

A.4 Robustness to the Measurement of Reform Intensity

	REFORM INTENSITY†			REFORM INTENSITY‡		
	Ratio Never Enrolled	Average Education - All -	Average Education - Enrollees -	Ratio Never Enrolled	Average Education - All -	Average Education - Enrollees -
	(1)	(2)	(3)	(1)	(2)	(3)
Reform intensity (years)	-0.02***	0.11***	0.03*	-0.03***	0.11**	0.02
	(0.00)	(0.03)	(0.02)	(0.01)	(0.04)	(0.02)
Reform × Female	-0.02***	0.19***	0.04***	-0.02***	0.18***	0.04***
	(0.00)	(0.01)	(0.00)	(0.00)	(0.02)	(0.01)
Reform × Catholic	0.00	0.00	0.02*	0.00	0.01	0.02
	(0.00)	(0.02)	(0.01)	(0.00)	(0.02)	(0.01)
Reform × Female× Catholic	-0.01***	0.04***	-0.00	-0.01***	0.05***	-0.00
	(0.00)	(0.02)	(0.00)	(0.00)	(0.02)	(0.01)
Reform × Protestant	0.00	-0.01	-0.00	0.00	-0.01	-0.00
	(0.00)	(0.03)	(0.01)	(0.01)	(0.03)	(0.01)
Reform × Female × Protestant	-0.0048***	0.03***	0.00	-0.01**	0.04**	0.00
	(0.00)	(0.01)	(0.00)	(0.00)	(0.02)	(0.00)
Reform × Village authority	-0.00	0.00	-0.01	-0.00	0.01	-0.01
	(0.00)	(0.01)	(0.01)	(0.00)	(0.02)	(0.01)
Reform × Female × Village	-0.00	0.01	0.01**	-0.00	0.01	0.01**
	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)
Adjusted within R^2	0.10	0.10	0.03	0.10	0.10	0.03
F	74	68	18	59	65	19
Nr. clusters	103	103	103	103	103	103
N	7,004	7,004	7,004	7,004	7,004	7,004

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Errors are clustered at the level of districts and are presented in parenthesis. All regressions are based on the preferred specification, which includes the following fixed effects: District, Gender, Age, Period, District × Period, Age × Period, Age × District. The Catholic, Protestant and village-authority variables are standardised values of the normalised number of schools per 100,000 children aged 7–13 in 2012; $\mu = 1$ and $\sigma = 1$. All regressions include the 14–30 age groups of 2002 and 2012. In terms of birth cohorts, this means: 1972–98. The 1989–94 cohorts have been partially treated, and the 1995–97 cohorts have been fully treated. Census 2002 data is used to gauge reform intensity.

† The intensity variable is computed based on the educational performance of individuals who were aged 19–25 at the time of the reform (7 birth-year cohorts, 1977–83).

‡ The intensity variable is computed based on the educational performance of individuals who were aged 26–32 at the time of the reform (7 birth-year cohorts, 1970–76). Reform intensity is district-gender-birth-cohort specific. The sample only includes the districts of Mainland Tanzania as of 1988.

A.5 Robustness to the Removal of Individuals as Region of Residence is Different from Region of Birth

Explanatory variables	RATIO OF NEVER-ENROLLED INDIVIDUALS - All individuals -					YEARS OF PRIMARY EDUCATION - All individuals -					YEARS OF PRIMARY EDUCATION - Only enrollees -				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Reform intensity (years)	-0.03***	-0.03***	-0.03***	-0.03***	-0.03***	0.12***	0.13***	0.13***	0.13***	0.14***	0.03	0.03*	0.03	0.03	0.04*
	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Reform × Female	-0.02***	-0.02***	-0.02***	-0.02***	-0.02***	0.17***	0.18***	0.17***	0.17***	0.18***	0.04***	0.04***	0.04***	0.04***	0.04***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Reform × Catholic		0.00			0.00		0.00			0.01		0.02			0.02*
		(0.00)			(0.00)		(0.02)			(0.02)		(0.01)			(0.01)
Reform × Female× Catholic		-0.01***			-0.01***		0.06***			0.05***		-0.00			-0.00
		(0.00)			(0.00)		(0.02)			(0.02)		(0.00)			(0.00)
Reform × Protestant			0.00		0.00			-0.01		-0.01			0.00		0.00
			(0.00)		(0.00)			(0.03)		(0.03)			(0.01)		(0.01)
Reform × Female × Protestant			-0.01***		-0.01***			0.04***		0.04***			0.00		0.00
			(0.00)		(0.00)			(0.02)		(0.01)			(0.00)		(0.00)
Reform × Village authority				-0.00	-0.00				0.00	0.00				-0.01	-0.01
				(0.00)	(0.00)				(0.01)	(0.01)				(0.01)	(0.01)
Reform × Female × Village				-0.00	-0.00				0.02	0.02				0.01**	0.01**
				(0.00)	(0.00)				(0.02)	(0.01)				(0.00)	(0.00)
Adjusted within R^2	0.09	0.10	0.09	0.09	0.10	0.09	0.09	0.09	0.09	0.10	0.03	0.03	0.02	0.03	0.03
F	259	135	120	152	71	223	134	109	131	70	49	29	25	26	15
Nr. clusters	103	103	103	103	103	103	103	103	103	103	103	103	103	103	103
N	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004	7,004

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Errors are clustered at the level of districts and are presented in parenthesis. Dependent variables have been computed based on the sample of individuals who have *not* changed their region of residence since birth. All regressions are based on the preferred specification, which includes the following fixed effects: District, Gender, Age, Period, District × Period, Age × Period, Age × District. The Catholic, Protestant and village-authority variables are standardised values of the normalised number of schools per 100,000 children aged 7–13 in 2012; $\mu = 1$ and $\sigma = 1$. All regressions include the 14–30 age groups of 2002 and 2012. In terms of birth cohorts, this means: 1972–98. The 1989–94 cohorts have been partially treated, and the 1995–97 cohorts have been fully treated. Census 2002 data is used to gauge reform intensity. Reform intensity is computed based on the educational performance of individuals who were aged 19–32 at the time of the reform (born 1970–83). The sample includes the districts of Mainland Tanzania as of 1988.

A.6 Railway Infrastructure in Tanganyika Cca. 1950
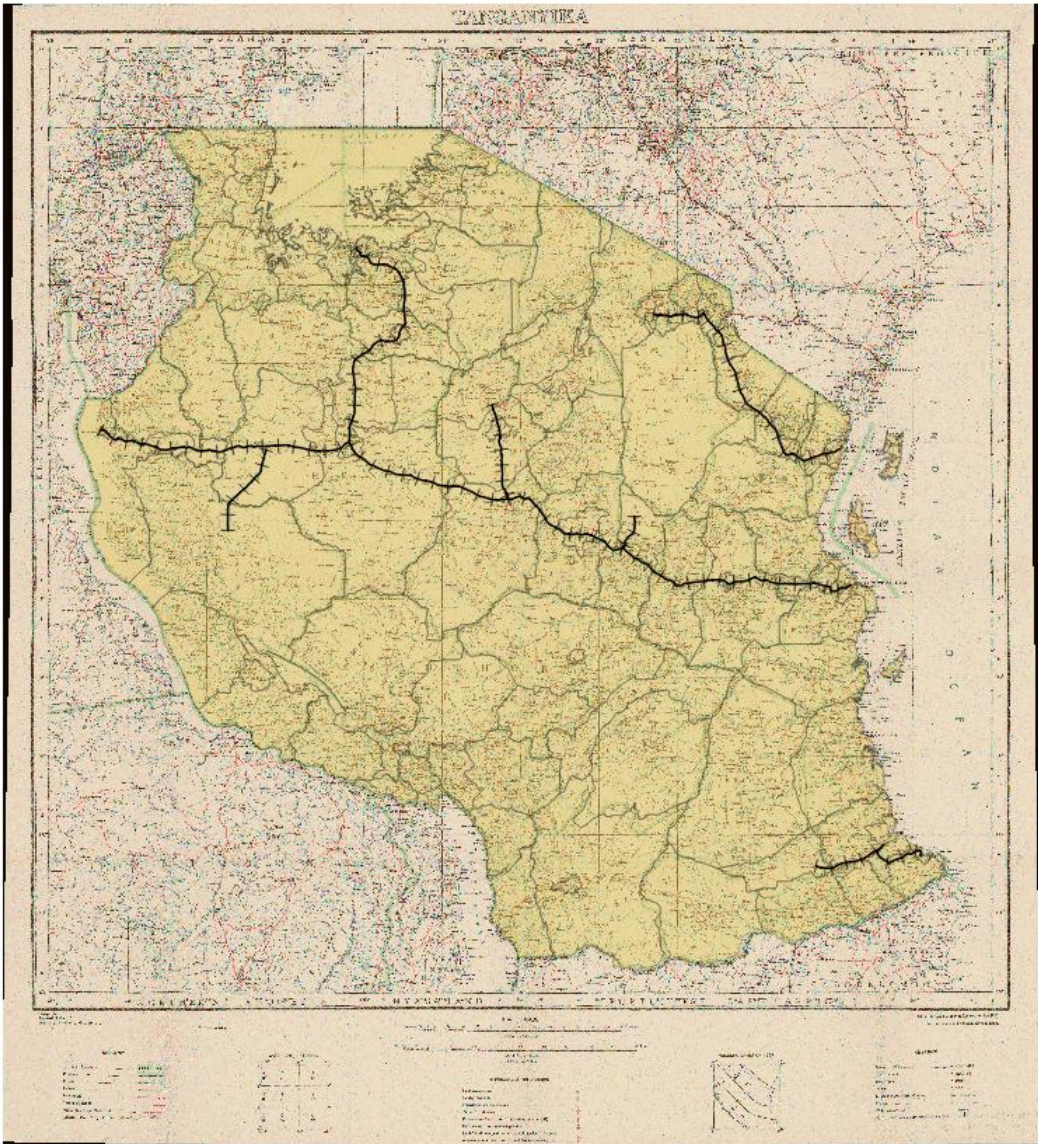
Source: Tanganyika. Department of Lands and Surveys. Hosted by the Princeton University Library and available online at https://catalog.princeton.edu/catalog/10159264.

References

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. (2017). When Should You Adjust Standard Errors for Clustering? *NBER Working Paper*(w24003).

Acemoglu, D., Johnson, S., & Robinson, J. (2005). The Rise of Europe: Atlantic Trade, Institutional Change, and Economic Growth. *American Economic Review*, *95*(3), 546–579.

Alesina, A., Giuliano, P., & Nunn, N. (2013). On the Origins of Gender Roles: Women and the Plough. *The Quarterly Journal of Economics*, *128*(2), 469–530.

Al-Samarrai, S., & Reilly, B. (2000). Urban and Rural Differences in Primary School Attendance: An Empirical Study for Tanzania. *Journal of African Economies*, *9*(4), 430–474.

Ashraf, N., Bau, N., Nunn, N., & Voena, A. (2020). Bride Price and Female Education. *Journal of Political Economy*, *128*(2), 591–641.

Becker, S. O., & Woessmann, L. (2008). Luther and the Girls: Religious Denomination and the Female Education Gap in Nineteenth-Century Prussia. *Scandinavian Journal of Economics*, *110*(4), 777–805.

Bleakley, H., & Lin, J. (2012). Portage and Path Dependence. *The Qarterly Journal of Economics*, *127*(2), 587–644.

Bruhn, M., & Gallego, F. A. (2012). Good, Bad, and Ugly Colonial Activities: Do they Matter for Economic Development? *Review of Economics and Statistics*, *94*(2), 433–461.

Buchert, L. (1991). *Politics, Development and Education in Tanzania 1919-1985: An Historical Interpretation of Social Change.* (Doctoral dissertation. Institute of Education, University of London.)

Calvi, R., & Mantovanelli, F. G. (2018). Long-Term Effects of Access to Health Care: Medical Missions in Colonial India. *Journal of Development Economics*, *135*, 285–303.

Cameron, J. (1967). The Integration of Education in Tanganyika. *Comparative Education Review*, *11*(1), 38–56.

Cameron, J., & Dodd, W. A. (1970). The Development of Education, 1919–1945. In *Society, Schools and Progress in Tanzania* (pp. 58–76). Pergamon.

Chachage, C. S. L. (1988). British Rule and African Civilization in Tanganyika. *Journal of Historical Sociology*, *1*(2), 199–223.

Chicoine, L. (2019). Schooling with Learning: The Effect of Free Primary Education and Mother Tongue Instruction Reforms in Ethiopia. *Economics of Education Review*, *69*, 94–107.

Chicoine, L. (2020). Free Primary Education, Fertility and Women's Access to the Labor Market: Evidence from Ethiopia. *Worl Bank Policy Research Working Paper, 9105*.

Cilliers, J., Mbiti, I. M., & Zeitlin, A. (2020). Can Public Rankings Improve School Performance? Evidence from a Nationwide Reform in Tanzania. *Journal of Human Resources*, 0119–9969R1.

Deininger, K. (2003). Does Cost of Schooling Affect Enrollment by the Poor? Universal Primary Education in Uganda. *Economics of Education Review*, *22*(3), 291–305.

Dell, M. (2010). The Persistent Effects of Peru's Mining Mita. *Econometrica*, *78*(6), 1863–1903.

Frankema, E. (2010). The Colonial Roots of Land Inequality: Geography, Factor Endowments, or Institutions? *The Economic History Review*, *63*(2), 418–451.

Galabawa, C. (1990). *Implementing Educational Policies in Tanzania.* (Tech. Rep.). (World Bank Discussion Papers No. 86. Africa Technical Department Series.)

Gillette, A. L. (1977). Beyond the Non-Formal Fashion: Towards Educational Revolution in Tanzania.
(Centre for International Education. University of Massachussets.)

Government of Tanzania. (2001). Primary Education Development Plan I (2002–2006).

Government of Tanzania. (2006). Primary Education Development Plan II (2007–2011).

Grogan, L. (2009). Universal Primary Education and School Entry in Uganda. *Journal of African Economies*, *18*(2), 183–211.

Heckman, J., & Robb, R. (1985). Using Longitudinal Data to Estimate Age, Period and Cohort Effects in Earnings Equations. In *Cohort Analysis in Social Research* (pp. 137–150). Springer.

Hoogeveen, J., & Rossi, M. (2013). Enrolment and Grade Attainment Following the Introduction of Free Primary Education in Tanzania. *Journal of African Economies*, *22*(3), 375–393.

Huillery, E. (2009). History Matters: The Long-Term Impact of Colonial Public Investments in French West Africa. *American Economic Journal: Applied Economics*, *1*(2), 176–215.

Ishumi, A. (2014). Voices in Development Struggles in the South: Experiences in Education in Tanzania, 1961-2011. In Z. Babaci-Wilhite (Ed.), *Giving Space to African Voices: Rights in Local Languages and Local Curriculum* (pp. 49–65). Sense Publishers, The Netherlands.

Jerve, A. M. (2006). *Exploring the Research-Policy Linkage: The Case of Reforms in Financing Primary Education in Tanzania.* Chr. Michelsen Institute. (Working paper nr. 2006:3)

Kondylis, F., & Manacorda, M. (2012). School Proximity and Child Labor Evidence From Rural Tanzania. *Journal of Human Resources*, *47*(1), 32–63.

Krueger, A. B., & Pischke, J. S. (1992). The Effect of Social Security on Labor Supply: A Cohort Analysis of the Notch Generation. *Journal of Labor Economics*, *10*(4), 412–437.

Lihwa, F., Johnstone, C. J., Thomas, M. A., & Krause, B. (2019). Remoteness as a Gendered Construct. *Development in Practice*, *29*(4), 501–513.

Lucas, A. M. (2010). Malaria Eradication and Educational Attainment: Evidence from Paraguay and Sri Lanka. *American Economic Journal: Applied Economics*, *2*(2), 46–71.

Lucas, A. M. (2013). The Impact of Malaria Eradication on Fertility. *Economic Development and Cultural Change*, *61*(3), 607–631.

Lucas, A. M., & Mbiti, I. M. (2012a). Access, Sorting and Achievement: The Short-Run Effects of Free Primary Education in Kenya. *American Economic Journal: Applied Economics*, *4*(4), 226–53.

Lucas, A. M., & Mbiti, I. M. (2012b). Does Free Primary Education Narrow Gender Differences in Schooling? Evidence from Kenya. *Journal of African Economies*, *21*(5), 691–722.

Machin, S., Marie, O., & Vujić, S. (2011). The Crime Reducing Effect of Education. *The Economic Journal*, *121*(552), 463–484.

Mason, A. D., & Khandker, S. R. (1996). *Measuring the Opportunity Costs of Children's Time in a Developing Country: Implications for Education Sector Analysis and Interventions* (Tech. Rep.). World Bank. (Report No. 16020)

Mbilinyi, M., Mbughuni, P., Meena, R., & Olekambaine, P. (1991). Equity Is Not Enough. Gendered Patterns of Education, Employment and Aid. In M. Mbilinyi & Mbughuni (Eds.), *Education in Tanzania with a Gender Perspective* (pp. 24–33). SIDA. (Education Division Documents No. 53)

Mbiti, I., Muralidharan, K., Romero, M., Schipper, Y., Manda, C., & Rajani, R. (2019). Inputs, Incentives, and Complementarities in Education: Experimental Evidence from Tanzania. *The Quarterly Journal of Economics*, *134*(3), 1627–1673.

Minnesota Population Center. (2018). *Integrated Public Use Microdata Series*. Minneapolis, MN: IPUMS. (Version 7.1 [dataset].)

Montgomery, M. (2017). Colonial Legacy of Gender Inequality: Christian Missionaries in German East Africa. *Politics & Society*, *45*(2), 225–268.

Muralidharan, K., & Prakash, N. (2017). Cycling to School: Increasing Secondary School Enrollment for Girls in India. *American Economic Journal: Applied Economics*, *9*(3), 321–50.

Naritomi, J., Soares, R. R., & Assunção, J. J. (2012). Institutional Development and Colonial Heritage within Brazil. *The Journal of Economic History*, 393–422.

National Bureau of Statistics. (2002). *Household Budget Survey 2000–01* (Tech. Rep.). (Key findings)

National Bureau of Statistics, & Office of Chief Government Statistician. (2013). *2012 Population and Housing Census: Population Distribution by Administrative Areas* (Tech. Rep.).

Nishimura, M., Yamano, T., & Sasaoka, Y. (2008). Impacts of the Universal Primary Education Policy on Educational Attainment and Private Costs in Rural Uganda. *International Journal of Educational Development*, *28*(2), 161–175.

Nunn, N. (2009). The Importance of History for Economic Development. *Annual Review of Economics*, *1*(1), 65–92.

Nunn, N. (2014a). Gender and Missionary Influence in Colonial Africa. *African Development in Historical Perspective*.

Nunn, N. (2014b). Historical Development. In *Handbook of Economic Growth* (Vol. 2, pp. 347–402). Elsevier.

Nunn, N., & Wantchekon, L. (2011). The Slave Trade and the Origins of Mistrust in Africa. *American Economic Review*, *101*(7), 3221–3252.

Oketch, M., & Rolleston, C. (2007). Policies on Free Primary and Secondary Education in East Africa: Retrospect and Prospect. *Review of Research in Education*, *31*(1), 131–158.

Olekambaine, P. (1991). Primary Education. In M. Mbilinyi & Mbughuni (Eds.), *Education in Tanzania with a Gender Perspective* (pp. 34–48). SIDA. (Education Division Documents No. 53)

Omari, I. M., Mbise, A., Mahenge, S., Malekela, G., & Besha, M. (1983). *Universal Primary Education in Tanzania*. IDRC, Ottawa, ON, CA.

Pop-Eleches, G. (2007). Historical Legacies and Post-Communist Regime Change. *The Journal of Politics*, *69*(4), 908–926.

Redding, S. J., Sturm, D. M., & Wolf, N. (2011). History and Industry Location: Evidence from German Airports. *Review of Economics and Statistics*, *93*(3), 814–831.

Siwale, E., & Sefu, M. (1977). The Development of Primary Education in Tanzania. (Brock University, St. Catharines, Ontario, Canada, ERIC No.: ED142280.)

Somerset, A. (2009). Universalising Primary Education in Kenya: The Elusive Goal. *Comparative Education*, *45*(2), 233–250.

Sumra, S. (2017). *The Impact of the Implementation of Fee-Free Education Policy on Basic Education in Tanzania: A Qualitative Study* (Tech. Rep.). HakiElimu.

Tabetah, J. A. (1982). *The Effects of Post-Independence Reform Policy on Public Education in Africa: The Case of Tanzania*. (Doctoral dissertation. Virginia Polytechnic Institute and State University.)

Trewartha, G. T., & Zelinsky, W. (1954). Population Patterns in Tropical Africa. *Annals of the Association of American Geographers*, *44*(2), 135–162.

Valente, C. (2019). Primary Education Expansion and Quality of Schooling. *Economics of Education Review*, *73*, 101913.

Vavrus, F., & Moshi, G. (2009). The Cost of a 'Free' Primary Education in Tanzania. *International Critical Childhood Policy Studies Journal*, *2*, 31-42.

Vogler, J. P. (2019). Imperial Rule, the Imposition of Bureaucratic Institutions, and Their Long-Term Legacies. *World Politics*, *71*(4), 806–863.

World Bank. (2001). *Report and Recommendation of the President of the International Development Association to the Executive Directors on a Proposed Adjustment Credit to the United Republic of Tanzania for a Primary Education Development Programme Project.* (Report nr. P7466 TA)

World Bank. (2005). *Implementation Completion Report for the Primary Education Development Programme.* (Report nr. 32071)

World Bank, & UNICEF. (2009). *Abolishing School Fees in Africa: Lessons from Ethiopia, Ghana, Kenya, Malawi, and Mozambique.* World Bank.

Zenebe Gebre, T. (2019). Free Primary Education, Timing of Fertility, and Total Fertility. *The World Bank Economic Review*.