

Colonial Institutions, Marriage Markets, and HIV: Evidence from Mozambique^{*}

Jon Denton-Schneider[†]

University of Michigan

August 15, 2021

Preliminary and incomplete. Please do not cite or circulate without permission.

Abstract

This paper links Africa's history to its HIV epidemic through colonial institutions' lasting effects on marriage markets. I exploit the arbitrary border within Mozambique between two regimes common across the continent: one that pushed over 50,000 young men annually into temporary labor migration (1897-1965) and another that subjected them to forced labor (1891-1942). Historians contend the migrant-sending institution fundamentally altered marriage markets in that region. Using colonial census data, I show that young men there still married earlier and were closer in age to their wives two decades after the end of the forced labor institution, even though migration rates had converged. Because smaller age disparities reduce HIV risk, I examine seroprevalence today and find it is nearly 50 percent (10 p.p.) lower in the former migrant-sending region. The data suggest that age disparities and associated behaviors are the main channel for this effect.

*I am indebted to Hoyt Bleakley and Eduardo Montero for their thoughtful guidance and extensive patience. I thank them and Ach Adhvaryu, Elisa Maffioli, Ana Reynoso, Dean Yang, and seminar participants at Michigan and DemSemX for excellent comments and feedback. I also thank Zijun Li, Abby Lin, and Emily Schmitt for providing outstanding research assistance, and Carlos Lopes Brito for sharing a copy of the 1940 map of Mozambique. I am grateful for financial support from a National Institute on Aging training grant (T32AG000221) and the Marshall Weinberg Endowment, both through the University of Michigan Population Studies Center.

[†]Department of Economics and Population Studies Center. Email: jdds@umich.edu

1. Introduction

After European powers carved up Africa at the end of the nineteenth century, they established institutions to extract wealth from and govern their colonies. [Amin \(1972\)](#) underscored the significance of these regimes in African development by using them to group colonies into "macro-regions" where one was predominant. In East and Southern Africa, the main institution was the *labor reserve* that provided mines and settler farms with an "army of short-term male [migrant] labor," while in the Congo Basin it was the *concession*, or a grant of "land (and the Africans living on [it]) to private companies" ([Roberts, 2017](#), p. 585).¹ Figure 1, a map of the present-day countries where these two were the primary regimes, shows how pervasive they were in these regions.

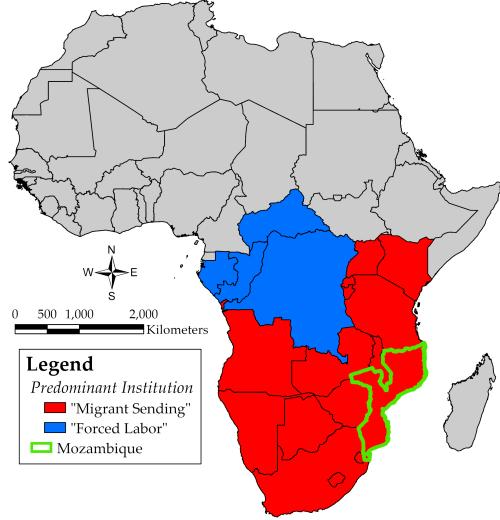
Because these migrant-sending and forced labor institutions organized much of the continent's economic activity during the colonial era, they may have meaningfully shaped the historical trajectories of the regions that were under them—today, they are some of the world's poorest or have its highest rates of HIV prevalence. Thus it is important to understand whether the choice of regime shaped such outcomes, and if so, through which channels.

However, as yet there is no causal evidence on the comparative impacts of historical assignment to one of these two institutions. This question is important because a colonial power's choice of regime in what is now the developing world was almost never between imposing an inclusive or extractive one (using the [Acemoglu and Robinson \(2012\)](#) terminology), but rather between different forms of extraction. Therefore, evaluating the comparative impacts of these common regimes can sharpen our understanding of how institutions influence development outcomes today.

The main challenge in generating such evidence is that institutions were not

¹ A third institution—the *colonial trade economy*, under which coerced peasant farmers produced cash crops for export—predominated in West Africa. Because it combined features of migrant-sending and forced labor, I focus in this paper on the sharper distinctions between these two regimes.

Figure 1: Institutional Macro-Regions in Africa



Notes: Map groups present-day countries by their [Amin \(1972\)](#) predominant colonial institution: the labor reserve ("migrant-sending") or the concession ("forced labor"). Mozambique is the southeastern country outlined in green.

randomly assigned. Instead, the distinct human and natural resource geographies of a macro-region led the various European powers with colonies in it to rely primarily on a single extractive regime. This relationship complicates simple comparisons between them, and such factors (e.g., climate, crop suitability) are also likely to directly affect many outcomes of interest.

The point of departure for this paper is that while an institution may have predominated in a macro-region, there were exceptional colonies like Mozambique in which it was not the only one ([Alexopoulou and Juif, 2017](#)). Located in the labor reserve region in southeastern Africa, this former Portuguese colony contained both a migrant-sending and forced labor institution. And importantly for causal inference, the border between two of them was arbitrary, consisting almost entirely of straight lines defined by latitude and longitude.

In this paper, I exploit this arbitrary border between one of Africa's most important migrant-sending institutions (1897-1965), which sent over 50,000 short-term labor migrants (henceforth circular migrants) to South African gold mines each year, and its

longest-lasting forced labor regime (1891-1942). Section 2 provides greater detail regarding these institutions' histories and the differences between them, but the main distinction according to historians was in men's labor mobility. While it was extreme to the south of the border in the migrant-sending region, the forced labor regime heavily restricted it so as to preserve its pool of conscript labor.

This arbitrary border allows me to estimate the causal impact of historical assignment to the former instead of the latter with a geographic regression discontinuity (RD) design, which I describe in Section 3. Using newly-digitized data from two colonial censuses of Mozambique, Section 4 examines the outcomes that were different along the border two years prior to the end of the forced labor institution (1940) and nearly two decades afterward (1960). I find that the migrant-sending regime had significantly higher rates of men's circular migration in 1940 but there were no longer any differences by 1960, showing the importance of the forced labor institution's mobility restrictions.

However, in spite of this convergence, marriage market outcomes remained different: in 1960, men in the former migrant-sending region married significantly earlier and age disparities between spouses were smaller, which suggests that circular migration had lasting impacts on marriage markets. I also examine schooling given the links between migration and human capital investment (e.g., [Yang, 2008](#); [Dinkelman and Mariotti, 2016](#)). In contrast to these studies, I find lower rates of schooling for boys in the migrant-sending area in 1940, though this result is most likely due to changes in mission-provided education at this time rather than the effects of migration. By two decades later, I show that schooling had converged along with circular migration.

I turn in Section 5 to analyzing how these colonial-era differences (or lack thereof) would affect two of the most important outcomes in sub-Saharan Africa today: HIV prevalence and economic development. I argue that they should lead to lower seroprevalence in the former migrant-sending region due to the importance of age gaps between partners for the spread of the virus ([de Oliveira et al., 2017](#); [Schaefer et al.,](#)

2017). I then describe why we should expect differences in age disparities to have continued: the decades-long shock to men's marriageability in the migrant-sending region could have persistently reduced the pool of unmarried young men common in polygynous societies, and it could have shifted preferences for partners (Irons, 1983; Leclerc-Madlala, 2008; Henrich, Boyd and Richerson, 2012). Conversely, the convergence in migration and human capital accumulation in the colonial era imply no differences in development today.

The results in Section 6 are consistent with these hypotheses. I examine georeferenced blood test data from two waves of the Demographic and Health Surveys (DHS) in Mozambique and find a decrease of almost 50 percent (10 percentage points, or p.p.) in HIV prevalence just inside the former migrant-sending region. In contrast, I find no differences at the border today in wealth, schooling, or children's health.

Finally, I investigate in Section 7 whether present-day marriage and dating market outcomes change at the border as they did in the colonial period. I find that there still are smaller age disparities between spouses and sexual partners in the former migrant-sending region, and that behaviors associated with these disparities are less common there as well (Evans et al., 2019; Mabaso et al., 2021). Because I find no evidence to support a number of other potential explanations (e.g., genital ulcers, transactional sex, women's autonomy), these findings suggest that the long-lasting effects of these institutions is the main channel for the HIV result.

This paper to several literatures. Influential studies examine the effects of colonial institutions on modern outcomes (Acemoglu, Johnson and Robinson, 2001; Banerjee and Iyer, 2005; Dell, 2010; Michalopoulos and Papaioannou, 2014, 2016), and there is recent evidence specifically on concessions (Dell and Olken, 2020; Lowes and Montero, 2021a; Méndez-Chacón and Van Patten, 2021). However, we know little about the short- or long-run impacts of the choice of extractive institution, and about labor reserves in particular. There is also limited evidence on the long-run effects of circular migration (Dinkelman

and Mariotti, 2016; Khanna, Theoharides and Yang, 2020), even though it was “one of the most distinctive features of [the African] continent’s development” (Stichter, 1985, p. 1). In addition, this study adds to our understanding of the relationship between marriage markets and human capital (Chiappori, Iyigun and Weiss, 2009), especially health (Anderson, 2018). Finally, these results add to the growing literature on history as a determinant of health (Alsan and Wanamaker, 2018; Lowes and Montero, 2021b), especially the spatial distribution of HIV (Bertocchi and Dimico, 2019; Dwyer-Lindgren et al., 2019; Cagé and Rueda, 2020).

2. History of the Migrant-Sending and Forced Labor Institutions

The Berlin Conference of 1884-85 established effective occupation as the principle for European powers to maintain claims to their African colonies. To meet this standard in Mozambique, Portugal pursued a two-part strategy that was the most expedient given its financial crisis at the time and British and German plans to partition the territory should it fail (Newitt, 1995). The first component was to project the colonial state outward from port cities established in the sixteenth century into the surrounding regions, and the second part was to grant vast, mostly unexplored areas to private companies as concessions (Smith and Smith, 1985).

2.1. Assignment of Territory to Government or Company Rule

Portuguese officials initially used major rivers to define the borders between areas under colonial state and company rule. Leveraging its presence in Lourenço Marques (present-day Maputo) and on Mozambique Island, the government assigned to itself the areas from the southern international border to the Sabi River and around the then-capital between the Ligonha and Lurio Rivers. But it could not quickly establish state capacity between the Sabi River and (tributaries of) the Zambezi, or between the Lurio River and

the northern international border ([Newitt, 1995](#)).²

Therefore, in 1891 Portugal granted royal charters to the Mozambique Company to administer the former region and to the Niassa Company to administer the latter.³ However, the boundaries of the Mozambique Company's territory changed twice in the following decade. Figure 2 shows the final institutional boundaries. First, a royal decree in 1893 expanded the concession southward to a border defined almost entirely by latitude and longitude, citing the need to effectively occupy more of the colony:

Whereas the Mozambique Company has at its disposal important means of action, and consequently it is highly expedient that [lands south of the Sabi River] should be administered by that Company, so as to insure the proper development and defence [*sic*] of those territories; ... The administration and "exploitation" of the territory bounded ... [by the Sabi River, the Limpopo River] as far as the point where it is intersected by the 32nd meridian, ... the direct line starting from the last-named point as far as that where the 32nd meridian intersects the 22nd parallel of latitude, and [the line] following the course of the said parallel of latitude as far as the sea ... is granted to the Mozambique Company. ([Great Britain Foreign Office, 1901](#), pp. 601-602)

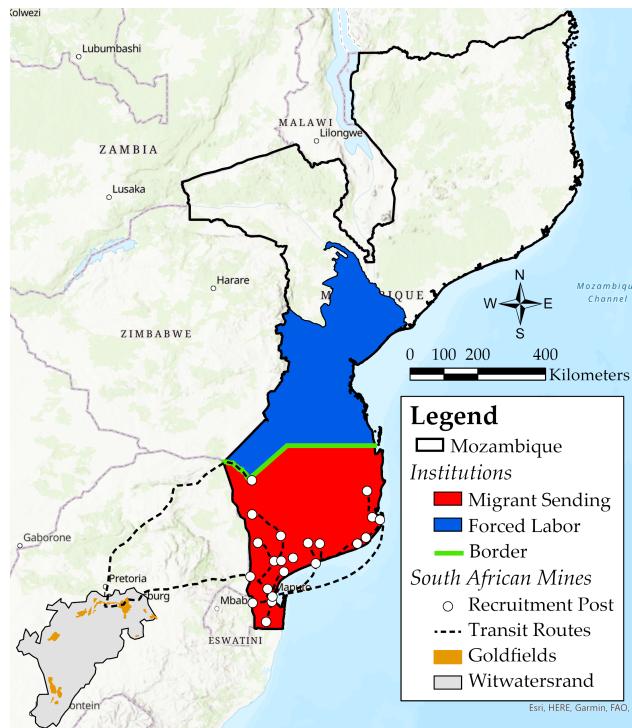
In contrast, the second border change resulted from the company losing the Barue District in the northwest of its territory in 1902.⁴ The local population had proved difficult to subdue and the suppression of a rebellion in that year required the assistance of Portuguese soldiers. Because the Mozambique Company did not repay the costs of the intervention, the state annexed Barue and incorporated it into the newly-created Tete Province that was under its administration ([Dokkum, 2020](#)).

² The area from the Zambezi basin to the Ligonha River featured a unique blend of African and Portuguese institutions known as *prazos* that emerged in the sixteenth century. Because they did not exist anywhere else in Africa and Portugal allowed them to maintain their modes of governance until the 1930s, I do not examine them in this paper. See [Newitt \(1969\)](#) for more on the *prazo* system.

³ The charters were originally for 25 years but both of their terms were lengthened a few years later: the Niassa Company's through 1929 and the Mozambique Company's until 1942. This extension made the Mozambique Company's chartered concession the only one in Africa to last beyond the 1920s ([Vail, 1976](#)). Other companies leased concessions in Mozambique (e.g., the Zambezia Company), but they did not have the same rights of governance or responsibility for development that a royal charter conveyed.

⁴ For consistency throughout the paper, I use the current administrative naming conventions: first-level units are provinces, second-level units are districts, and third-level units are administrative posts.

Figure 2: Migrant-Sending and Forced Labor Institutions in Southern Mozambique



Notes: Map shows the two institutions as well as the Witwatersrand goldfields and WNLA's recruitment station and transportation network from [Transvaal Chamber of Mines \(1946\)](#).

2.2. Choice and Establishment of Extractive Institution

With Mozambique effectively occupied, “extracting wealth from African peasant society became the principal objective” of government and company officials ([Newitt, 1995](#), p. 406). They quickly discovered that labor was the easiest resource to exploit and established different institutions to accomplish their respective goals.

2.2.1. Migrant-Sending Institution

The colonial state set up a migrant-sending institution in its territory to profit from preexisting labor flows across the border to the then-Transvaal Republic (the northeast of present-day South Africa).⁵ The 1886 discovery of the world’s largest gold deposits

⁵ Men could be absent from southern Mozambique for extended periods because “the role of the male in [these ethnic groups’] agricultural life was negligible” given that the savannah required little clearing and women could cultivate the loose soil ([Rita-Ferreira, 1960](#), p. 144). [Junod \(1912\)](#) and [Harris \(1959\)](#) also noted this phenomenon and the labor mobility it had historically permitted men.

on the Witwatersrand led to intense demand for African workers that men from Mozambique were vital in filling (Clarence-Smith, 1985).⁶ To keep wages low, the mining companies formed the monopsonistic Witwatersrand Native Labour Association (WNLA) to recruit workers on their behalf.

WNLA and Portuguese authorities signed several agreements beginning in 1897 that formalized recruitment in the colonial state's southern territory. The government derived revenues from all parts of this process: licensing fees from recruiters, payments from WNLA for each worker, and permit fees from each worker allowing them to work abroad. Portuguese officials in Johannesburg also taxed wages paid on the Witwatersrand to Mozambicans (Newitt, 1995).

These agreements also regulated miners' contracts and how they were paid. Contract durations were limited to one year with a possible six-month extension and a mandated rest period of six months back in Mozambique. In 1928, the colonial state and the South African government established deferred payment for miners by which they would receive half of their wages only after returning home (Wilson, 1972).⁷ The Portuguese had long argued for this provision because miners spent much of their wages on the Witwatersrand—often to buy status goods—rather than in Mozambique (Harries, 1994).

In return, the colonial state granted a monopoly on labor recruitment in its territory to WNLA, which also benefitted from Portugal's 1899 colonial labor code. The law pushed men ages 14 to 60 into wage labor by subjecting them "to the moral and legal obligation to seek to acquire through employment the means to subsist and improve their social condition" or face forced labor (Portugal, 1900, p. 647). To capitalize on its monopoly and the masses of men seeking paid employment, WNLA established a series of stations across southern Mozambique for recruiting workers as well as

⁶ Geologists estimate that one-third of all gold ever mined is from the Witwatersrand (Frimmel, 2019).

⁷ In addition, the Transvaal government agreed to send up to one-half of its rail traffic through Lourenço Marques, helping Portugal realize its ambition for the city to become a major port (Clarence-Smith, 1985). Because these migrant labor and freight flows contributed heavily to the colonial state's finances, to better manage them it moved its capital from Mozambique Island to Lourenço Marques in 1902 (Newitt, 1995).

transportation infrastructure to move them from there to the Witwatersrand. Figure 2 shows this network in 1946 and Appendix A1 shows that except for the depths of the Great Depression, from 1920 to 1942 between 50,000 and 75,000 men annually arrived at the Witwatersrand mines from southern Mozambique.

2.2.2. Forced Labor Institution

In its territory, the Mozambique Company established a forced labor institution to attract large companies and settler farmers with a consistent supply of cheap workers. It issued regulations in 1900 requiring the population in its territory to engage in six months of wage labor each year, though administrators often conscripted workers on behalf of local employers offering wages too low or working conditions too harsh. Ten years later, the company formalized this forced labor system by establishing a department that could use violence to round up the workers that employers demanded ([Guthrie, 2018](#)).

The forced labor bureaucracy conscripted tens of thousands of workers each year by using its police to reinforce the efforts of traditional authorities. According to correspondence between company administrators, it was common for them to tell chiefs “that on such and such a date they had to supply a certain number of men to go work; generally, . . . because [some] cannot manage to organize the number of workers requested, one or more police go to help the chiefs who fell short” (as cited in [Allina, 2012](#), p. 50). Another method of ensuring compliance was to punish wives of men who tried to flee the forced labor system ([Guthrie, 2018](#)). The company also dissuaded many from attempting to work abroad by impressing “workers returning from abroad . . . into forced labor almost immediately, such that they . . . could not go home for any length of time unless they were willing to [be conscripted]” ([Allina, 2012](#), p. 58).

The company abolished its forced labor bureaucracy in 1926 as a response to a League of Nations report on labor practices in Portuguese colonies, which noted that “the blacks here [in the concession] tell the planters that they are the slaves of the Mozambique

Company" ([Ross, 1925](#), p. 53). However, employers soon complained that they could not find enough workers without the forced labor system. To push men into returning to these jobs, in 1927 the company doubled the annual hut tax so they would have to find wage labor and mandated that males over age 14 carry a pass book containing their picture, work history, tax payments, and place of residence. Officials frequently conducted sweeps checking that men had their pass books and met the six-month work requirement—the punishment for noncompliance was forced labor ([Allina, 2012](#)).

2.3. Narrative Comparisons of the Institutions

Given the rampant extraction of wealth from labor under both institutions, albeit in different forms, [Allina \(2012, p. 94\)](#) contended that "the [migrant-sending region] was governed by the Portuguese colonial state no less exploitatively than [the forced labor region was] by the company itself, and under the same labor code, if with a thinner presence on the ground." Similarly, [Harries \(1994, p. 175\)](#) argued that "Portugal was the chief recipient of the profits of [circular migration, which] . . . held back the development of southern Mozambique" despite the wages earned on the Witwatersrand.

Nonetheless, there may have been important differences between the two institutions in marriage outcomes as a result of circular migration. Historians have closely linked the two, arguing that in Southern African societies practicing brideprice (or *lobola*), "one of the primary reasons that men took up migrant labor was to obtain the money necessary for paying bridewealth. . . . Since most men intended to marry in their home areas, [it also] was critical in . . . persuading them to return home" ([Guthrie, 2018, p. 72](#)). Both [Junod \(1912\)](#) and [Fuller \(1955\)](#) noted that young men worked in the mines once or twice prior to marriage, implying many stopped migrating after completing their *lobola*.

Another contrast between the two regions was in who provided schooling to Africans, though it was not available to the vast majority of children in either one. While in the migrant-sending region Protestant missions established village schools and there were

some state-run rudimentary schools in densely populated areas, the company actually supported Catholic mission schools in its territory ([Allina, 2012](#); [Morier-Genoud, 2019](#)).⁸ Following the colonial state's closure of many of its village schools in 1930 due to concern over foreign and Protestant influences on the population, Catholic missions began to fill the gap but it is unclear how quickly they were able to do so ([Helgesson, 1994](#)).⁹

2.4. After the End of the Forced Labor Institution

The 1928 rise to power of the Portuguese autocrat Salazar brought the end of company rule in Mozambique. Fixated on national sovereignty, especially in the colonies, Salazar's government banned the granting of concessions in 1930, making it clear the Mozambique Company's charter would expire twelve years later ([Newitt, 1995](#)). In line with this focus, Appendix [A2](#) provides examples of the rhetoric surrounding the region's 1942 handover to the colonial state. Accompanying it was a territorial reorganization: the 1943 map in Appendix [A3](#) shows the erasure of the Mozambique Company's southern boundary as the provincial border moved back to the Sabi River.

The end of the forced labor institution implied an end to its restrictions on men's labor mobility. The visible rise in Appendix [A1](#) in the number of Mozambicans on the Witwatersrand after 1942 is consistent with several thousand men from the forced labor region joining the circular migration flows each year. The Salazar regime also unified education policy across Mozambique at this time, having the Catholic Church take over—and greatly increase—schooling for Africans ([Fernández Cebrián, 2021](#)).

The extraction of wealth from labor continued until 1961 and in many areas shifted focus to forced cotton cultivation, which historians linked to support for the guerrilla independence movement ([Isaacman et al., 1980](#); [Guthrie, 2016](#)). It succeeded in ending

⁸ A significant number of mine workers joined Protestant churches while on the Witwatersrand and missionaries followed them back to southern Mozambique, where they established a presence that included educating Africans in their local languages ([Newitt, 1995](#)).

⁹ Helgesson (1994) noted that between 1929 and 1930, the number of Methodist village schools fell from 200 to six and their student population fell from over 5,400 to under 700.

Portuguese rule in 1975 after a ten-year struggle, though the newly-independent country quickly fell into turmoil. To further destabilize it, apartheid-era South Africa sharply cut the number of Mozambicans on the Witwatersrand (see Appendix A1) and its security services aided the violent rebels in Mozambique's 1977-92 civil war (Weinstein, 2006). The country became one of the world's poorest during this period, and shortly after stability returned its HIV epidemic finally began to explode (Audet et al., 2010).¹⁰

3. Colonial Data and Empirical Strategy

3.1. Data

To compare the impacts of the migrant-sending and forced labor institutions while Mozambique was still under Portuguese rule, I digitized summaries of the colony's 1940 and 1960 censuses by district (Repartição Nacional de Estatística, 1942; Direcção Provincial dos Serviços de Estatística, 1966). Appendix A4 contains examples of district-level summary tables from each census.

The 1940 data are the best available regarding the populations living under the two institutions while they both still existed. This census occurred two years before the end of the Mozambique Company's forced labor regime and it was the first one in the colony's history that met basic standards for accuracy (Darch, 1983; Harrison, 1998; Havik, 2013). Similarly, the 1960 data allow for the most reliable and longest-run comparison of the two regions during the colonial period. This census took place 18 years after the forced labor institution ended and it was the last one before the start of the Mozambican War of Independence (1964-74). As such, it does not suffer from the data collection problems that can arise when governments participate in internal conflicts (Barakat et al., 2002).

¹⁰ The HIV epidemics in Mozambique and Namibia—whose decades-long civil war ended in 1990—were in the exponential growth phase in the late 1990s while those in other Southern African countries had already matured. The implication is that Mozambique's and Namibia's began substantially later, likely because internal conflict limited mobility and thus the transmission of the virus (Iliffe, 2006).

3.1.1. Outcomes of Interest

I focus on outcomes in three domains: labor markets, marriage markets, and human capital accumulation. The first two are of interest because of the historical narratives in Section 2, which emphasize men’s circular migration and its effects on women’s work and marriage. I also include the third domain because of its relationship to economic development and because enumerators in the colonial period could measure it far more easily and accurately than income or wealth.

The labor market variables of interest are the share of males aged 15 to 64 (“prime-aged men”) who were circular migrants and the share of prime-aged women in agricultural occupations.¹¹ To examine differences in marriage market outcomes, I compute the ratio of ever-married men to ever-married women within a 10-year age group. For human capital accumulation, the outcome of interest is the share of boys and girls ages 5 to 14 who were enrolled in school at the time of enumeration.¹²

3.1.2. Georeferenced Sample

I match these district-level data to administrative maps of Mozambique from each year ([Saldanha, 1940](#); [Ministério do Ultramar, 1959](#)). Figure 3 shows district boundaries and centroids in the areas under the two institutions. I restrict the sample to districts within the two provinces south of the institutional border and the one north of it, and exclude the two major cities when the census summaries report their data separately.¹³

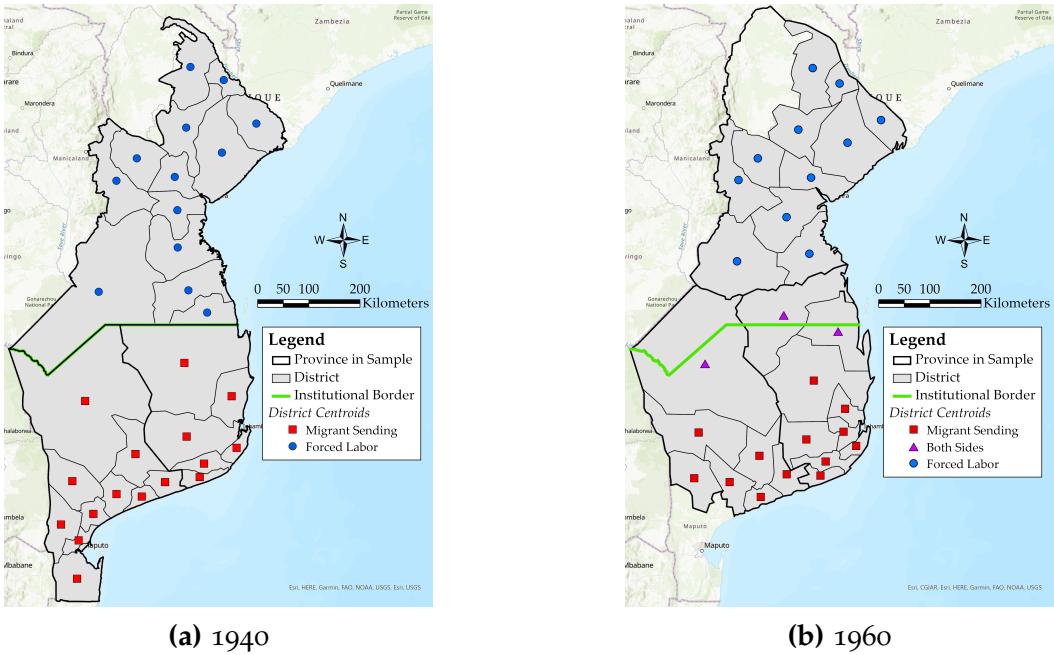
While 1940 boundaries respected the institutional border, districts after the 1942 territorial reorganization did not. For the three with area on both sides in 1960—the only ones whose centroids were within 100 km of the border—I assign them to the

¹¹ Each census’s questionnaire asked whether a man worked abroad but the 1960 summary tables grouped circular migrants into a category with all men who worked in a mine regardless of its location. However, nearly all men in this category worked abroad rather than in the few mines in Mozambique, so I consider this variable a measure of circular migration.

¹² The numerator excludes those who had left school before enumeration, so it is weakly less than the fraction of children who had ever attended school.

¹³ These two cities are Lourenço Marques and Beira, the provincial capital in the forced labor region.

Figure 3: Maps of Georeferenced Colonial Census Data



Notes: Maps show the districts in each institution matched to census data and their centroids.

institution containing their centroids. Below, I discuss how doing so affects the results.

3.2. Empirical Strategy

To compare the impacts of the two institutions during the colonial era, I estimate the following RD specification:

$$y_d = \alpha + \tau \text{MigrantSending}_d + f(\text{Distance}_d) + \text{Lon}_d + \epsilon_d \quad \text{for } d \in B \quad (1)$$

where y_d is the outcome of interest for district d in the set B defined by the bandwidth restrictions above.¹⁴ The explanatory variables are MigrantSending_d , an indicator for whether d 's centroid is in that institution; $f(\text{Distance}_d)$, the RD polynomial controlling for smooth functions of a centroid's distance to the institutional border; and Lon_d , a centroid's longitude coordinate, which Kelly (2020) recommends including in RD designs to capture east-west trends.¹⁵ I use a local linear specification estimated

¹⁴ There are too few centroids near the border to estimate the Calonico, Cattaneo and Titiunik (2014) mean squared error (MSE) optimal bandwidth for colonial-era outcomes.

¹⁵ Distance_d has a near-perfect correlation with latitude ($\rho > 0.99$), so it accounts for north-south trends.

separately on each side with a triangular kernel (Cattaneo, Idrobo and Titiunik, 2019; Gelman and Imbens, 2019). Because observations are district-level means, I weight them by the population in the denominator (e.g., the number of prime-aged men in d when the outcome is the share who were circular migrants).

The coefficient τ in Equation (1) identifies the effect of historical assignment to the migrant-sending institution *relative to historical assignment to the forced labor institution*. The motivating idea is that because the border between them was arbitrary, Portuguese colonial officials quasi-randomly allocated the territory around it to one of the two institutions. I examine the arbitrariness of the border in the next section.

3.2.1. Addressing Concerns with Estimation and Inference

One issue for estimating τ with the 1960 data is that there are districts with area on both sides of the border. To the extent that they group observations from one institution with those from the other, these districts will tend to obscure differences between the two and thus bias RD point estimates toward zero. I highlight them in the RD plots so the influence they have on the estimation is clear.

An important concern when conducting inference in geographic RD designs is positive spatial autocorrelation (Kelly, 2020). Due to the relatively small area under examination and the slow rate at which many outcomes change across space, estimated standard errors may be too small due to similarity among neighbors. Intuitively, assuming the statistical independence of observations would overstate the information each one adds to the estimation, leading to inflated precision.

I take two steps to address this potential problem. First, I calculate the Colella et al. (2020) standard errors allowing for arbitrary spatial correlation between observations within 100 km of each other, imposing a linear decay in relationships over this bandwidth (“Bartlett kernel”). I report these standard errors in addition to the usual heteroskedasticity-robust ones. Second, I diagnose positive spatial autocorrelation in

the residuals by computing the Moran (1950) I -statistic, which is the slope of the line in a weighted regression of neighbors' values on each unit's value.¹⁶ To generate the spatial weighting matrix, I set the bandwidth so that each district has at least one other neighbor, impose a Bartlett kernel, and row standardize the entries so that I is between -1 and 1.¹⁷ I report the difference of the observed and expected I -statistics as the measure of spatial autocorrelation along with I 's standard deviation.¹⁸

3.3. Balance on Precolonial and Geographic Traits

The assumption underlying the RD design is that all other relevant factors changed smoothly at the institutional border. To help rule out differences in precolonial characteristics, Figure 4 shows that the border is entirely within one Murdock (1959) ethnic homeland. Additionally, the neighboring ethnicities are all part of the same cultural group, suggesting that important behaviors and characteristics were not substantially different along the border at the time of assignment to an institution.

To test whether aspects of the geographic and disease environments changed along the border, I divide Mozambique into 0.25×0.25 degree cells—approximately 25 km \times 25 km in the study area—and estimate Equation (1) clustering standard errors by third-level administrative unit (“administrative post”).¹⁹ Consistent with the border’s arbitrary nature, Table 1 shows that estimated changes in these variables just inside the migrant-sending institution are small relative to mean values in the forced labor region, and standard errors might actually be too small given the number of clusters (Cameron, Gelbach and Miller, 2008) and high levels of positive spatial autocorrelation. Nonetheless, many of these measures (or their inputs) are interpolated across space, so

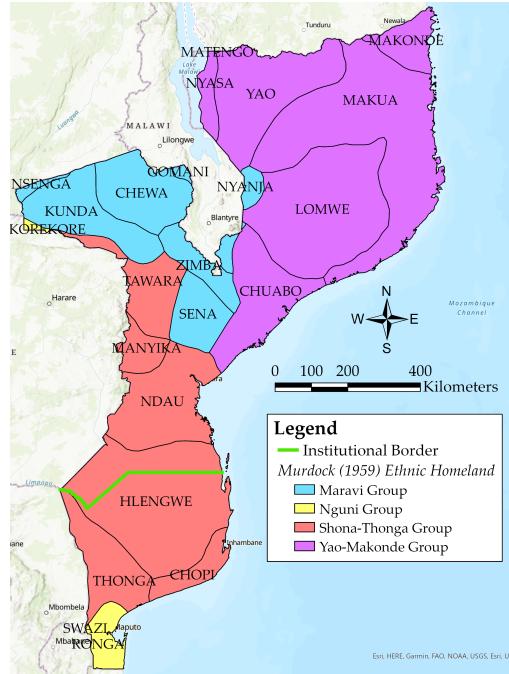
¹⁶ The I -statistic's expected value under no spatial autocorrelation is $\frac{-1}{N_s-1}$, which approaches zero from below as the number of unique sites (N_s) increases. An observed I greater (less) than the expectation indicates positive (negative) spatial autocorrelation, meaning neighboring sites have similar (dissimilar) values. Negative spatial autocorrelation implies that the effective degrees of freedom are *greater* than under statistical independence (Griffith and Arbia, 2010).

¹⁷ The resulting bandwidth is approximately 170 km for 1940 and approximately 140 km for 1960.

¹⁸ The asymptotic distribution of the I -statistic is standard normal.

¹⁹ See Section 5.3 for details on RD bandwidth selection when using geographically disaggregated data.

Figure 4: Ethnic Group Homelands in Mozambique



Notes: Map shows Murdock (1959) ethnic homelands by cultural group. The thick yellow line represents the concession-labor reserve border.

Table 1: Balance Tests at the Border

	Geographic Traits				Disease Suitability	
	Elevation	Rainfall	Slope	Soil Index	Malaria	TseTse
Migrant Sending	-1.798 (31.917) [23.254]	1.153 (6.533) [4.045]	0.014 (0.100) [0.054]	3.207 (3.887) [2.658]	-0.270 (0.368) [0.244]	-0.002 (0.009) [0.007]
Observations	167	105	144	115	139	173
Clusters	29	19	26	20	23	30
Bandwidth	131.5	79.4	120.6	94.3	107.0	144.3
Wild Cluster Bootstrap <i>p</i>	0.950	0.866	0.812	0.435	0.504	0.827
Spatial Autocorrelation	0.40	0.42	0.29	0.11	0.09	0.57
Spatial Autocorrel. SD	0.02	0.03	0.02	0.02	0.02	0.02
Forced Labor Mean	182.7	40.94	0.179	49.19	10.84	1.259

Notes: Observations are 0.25×0.25 degree cells. Standard errors clustered by administrative post are in parentheses. Regressions estimate a local linear RD specification on each side of the border using a triangular kernel and include longitude as a control. RD bandwidths are MSE-optimal (Calonico, Cattaneo and Titiunik, 2014). Variable definitions will be in an appendix.

failure to reject balance may provide only limited support for the RD assumption.

4. Effects of the Institutions during the Colonial Era

Table 2 reports the RD estimates for each outcome of interest two years before the abolition of the forced labor institution (Panel A) and 18 years after (Panel B). Figure 5 presents RD plots for selected outcomes in each year.

4.1. Labor Markets

Due to historians' emphasis on circular migration as the main difference between the institutions, I first examine whether it in fact changed at the border. Panel A Column (1) shows that prime-aged men just inside the migrant-sending institution were 21 p.p. more likely to be circular migrants. The effect size is very large given that just 5 percent of men in the forced labor region worked abroad. The RD plot in Figure 5a shows that while around 20 percent of men in two forced labor districts near the border were circular migrants, effectively none in the other districts were.

However, Panel B Column (1) shows convergence in men's circular migration after the forced labor institution's abolition as 16 percent of prime-aged men in that region worked in foreign mines.²⁰ This pattern implies that the forced labor regime significantly constrained men's mobility and its end in 1942 led to major changes in their occupational choices. Figure 5b shows that circular migration increased substantially inside the forced labor institution compared to the pre-abolition period.

Next, I examine whether differences between the institutions affected women's occupations. In spite of much higher rates of migration before the forced labor institution's abolition, Panel A Column (2) shows no difference at the border in the near-universal share of prime-aged women working in agriculture. There is also no

²⁰ This column excludes a district in the migrant-sending region that reported only one circular migrant, which is most likely an error in data reporting.

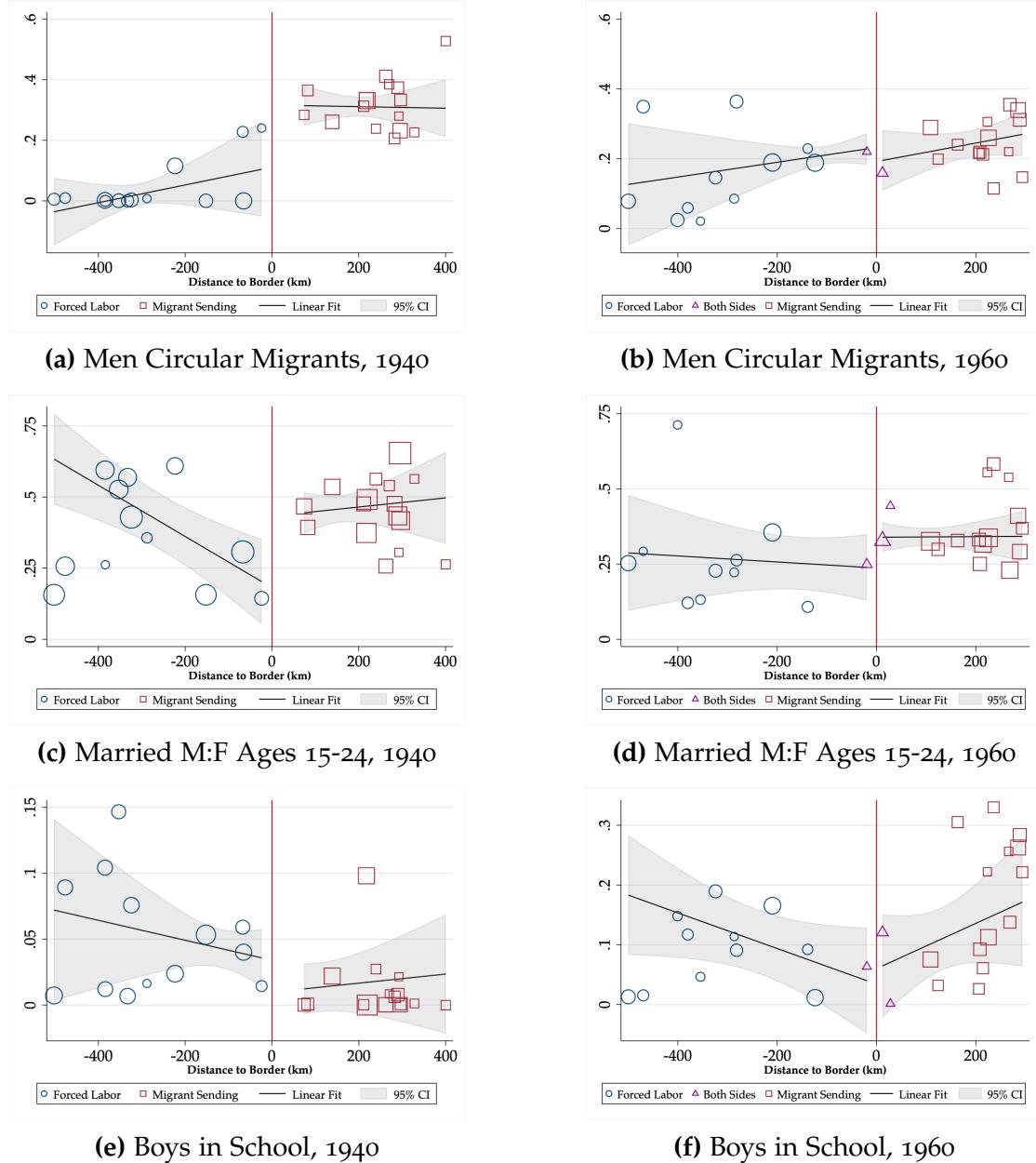
Table 2: Comparing Institutions during the Colonial Era

Panel A: 2 Years before Forced Labor Institution Abolished (1940)						
	Labor Markets		Marriage Markets		Human Capital	
	Men Circular Migrants (1)	Women in Agriculture (2)	Married M:F Ages 15-24 (3)	Married M:F Ages 25-34 (4)	Boys in School (5)	Girls in School (6)
Migrant Sending	0.207 (0.089) [0.085]	0.015 (0.018) [0.016]	0.269 (0.080) [0.061]	0.312 (0.113) [0.088]	-0.028 (0.016) [0.015]	-0.001 (0.006) [0.006]
Observations	29	29	28	28	29	29
Bandwidth	-503, 401	-503, 401	-503, 401	-503, 401	-503, 401	-503, 401
Spatial Autocorrel.	-0.14	-0.10	0.26	0.14	-0.03	-0.06
Spatial Autocor. SD	0.11	0.10	0.11	0.11	0.11	0.10
Forced Labor Mean	0.047	0.958	0.364	0.696	0.050	0.006

Panel B: 18 Years after Forced Labor Institution Abolished (1960)						
	Labor Markets		Marriage Markets		Human Capital	
	Men Circular Migrants (1)	Women in Agriculture (2)	Married M:F Ages 15-24 (3)	Married M:F Ages 25-34 (4)	Boys in School (5)	Girls in School (6)
Migrant Sending	-0.025 (0.049) [0.050]	0.006 (0.004) [0.004]	0.102 (0.057) [0.053]	0.106 (0.147) [0.131]	-0.014 (0.033) [0.029]	-0.026 (0.024) [0.021]
Observations	27	28	27	27	28	28
Bandwidth	-500, 294	-500, 294	-500, 294	-500, 294	-500, 294	-500, 294
Spatial Autocorrel.	-0.09	-0.02	-0.03	-0.15	0.44	0.24
Spatial Autocor. SD	0.12	0.11	0.12	0.10	0.12	0.13
Forced Labor Mean	0.163	0.997	0.267	0.635	0.089	0.041

Notes: Observations are districts. Colella et al. (2020) standard errors using a 100-km bandwidth and a Bartlett kernel are in parentheses. Regressions estimate a local linear RD specification on each side of the border using a triangular kernel and include longitude as a control. Data sources, variable definitions, and RD bandwidth selection criteria are in Section 3.1. Columns (3) and (4) in both panels and Panel B Column (1) each exclude an extreme outlier for that outcome as described in the text.

Figure 5: RD Plots for Colonial-Era Differences between Institutions



Notes: RD plots show the outcomes in each district. Local linear trends and 95% confidence intervals are estimated on each side of the institutional border using a triangular kernel and weighting by the relevant population. The running variable is distance in kilometers to the border. Plots for a variable have the same y-axis scale across years to facilitate comparisons. Data sources, variable definitions, and RD bandwidth selection criteria are in Section 3.1.

difference in Panel B Column (2). These results suggest that men's absences did not affect women's responsibility for food production. Indeed, this division of labor likely predated the institutions and enabled men's circular migration in the first place.

4.2. Marriage Markets

I then turn to comparing marriage outcomes across the institutional border, as narrative histories argue that earning bridewealth was an important motivation for men to work abroad. For brevity, I focus in Table 2 on the ratio of ever-married men to ever-married women for the 10-year age groups in which men engaged in circular migration (ages 15 to 24 and 25 to 34), as its marriage market impacts should have been most apparent for them. Appendix Table B1 contains results for the shares married by sex and age group.

Panel A Columns (3) and (4) show that while the two institutions still existed, the ratio of married men to women for both age groups increased around 30 p.p. just inside the migrant-sending region.²¹ These point estimates are large relative to the average ratios in the forced labor institution for ages 15 to 24 (just over 1:3) and 25 to 34 (2:3). Figure 5c shows the discontinuity for the younger age group, and the estimates in Appendix Table B1 Panel A imply that these differences were due entirely to young men's much higher rates of marriage in the migrant-sending region.

After the forced labor institution's abolition, the differences at the border in these ratios decreased but remained sizable. Panel B Columns (3) and (4) show that the size of the discontinuities fell to 11 p.p., though the estimate for ages 25 to 34 is imprecise. The RD plot in Figure 5d suggests that the decrease for ages 15 to 24 is due in part to districts with area on both sides of the border biasing the RD estimates toward zero. In spite of this effect, the estimates in Appendix Table B1 Panel B imply that young men in the migrant-sending region still married earlier than those across the border, while

²¹ These columns exclude a district in the forced labor institution that was an extreme outlier for this outcome: it had nearly equal numbers of married men and women ages 15 to 24, which is most likely an error in data reporting.

for the most part women on both sides continued to marry at similar rates. Notably, marriage on average was less common for men and women in both institutions across all age groups than it was two decades earlier.

4.3. Human Capital Accumulation

Finally, I examine whether differences between the institutions affected children's investment in human capital. Prior the forced labor institution's abolition—when Catholic missions were responsible for education there while the state was responsible for it in the migrant-sending region—Table 2 Panel A Column (5) shows that boys just inside the latter were 3 p.p. less likely to be in school. This effect size is large relative to the 5-percent enrollment rate for boys in the forced labor institution. Figure 5e shows this discontinuity visually.

However, after the forced labor institution ended and Catholic missions became responsible for education throughout Mozambique, the difference in boys' schooling at the border in Panel B Column (5) is smaller and less precise. Figure 5f shows that boys' enrollment rates increased markedly in most districts on either side of the border, and that the RD coefficient could be a result of the RD polynomial's slope. The former pattern is consistent with accounts of greater educational opportunities for Black Mozambican children after Catholic missions began providing basic schooling across the colony instead of just within the forced labor region ([Fernández Cebrián, 2021](#)).

In contrast to the pre-abolition result for boys, Panel A Column (6) shows that there were no differences in girls' schooling at the border and enrollment rates for girls in both institutions were effectively zero. But after the forced labor institution's abolition and the Catholic missions' takeover of basic education, girls' enrollment rates increased in most districts in the two regions. There is an imprecise negative coefficient in Panel B Column (6), though as for boys, it may simply be an artifact of the RD polynomial's slope (not shown).

5. Modern Data and Empirical Strategy

5.1. Conceptual Framework

The colonial-era patterns in the previous section—later marriages in the former concession with spouses likely farther apart in age, but no lasting differences in men’s circular migration or human capital accumulation—have clear implications for HIV prevalence and economic development today. With respect to the former, smaller age gaps between partners in sub-Saharan Africa lower the risk of contracting HIV ([Schaefer et al., 2017](#)). Intuitively, older men transmit HIV to younger women, who as they age transmit it to men of similar ages, continuing the cycle ([de Oliveira et al., 2017](#)). The result should be that where there are fewer age-disparate relationships, there is lower HIV prevalence and a later peak in women’s HIV age profile.

Therefore, if the marriage and dating patterns in Section 4 continue to today, all else equal they should reduce HIV prevalence in the former migrant-sending institution. One reason to expect they have is the additional decades of greater marriageability for young men. In most polygynous marriage markets in Southern Africa, many young women marry men of an older generation, reducing the pool of potential matches for young men. As they age, these single men will marry a young woman and perpetuate the cycle of age-disparate relationships. But in the former migrant-sending region, the long shock to young men’s marriageability may have dampened this cycle, allowing more young women to similarly-aged men and reducing the distortions in the next generation’s pool of potential matches.

Another reason to expect this pattern is one that Leclerc-Madlala offered when discussing marriage and dating in Southern Africa:

What we refer to now as age-disparate relationships . . . [has] antecedents in older practices that have long played a part in defining the nature of social life and the particular values and norms associated with sexuality. Many culturally inscribed assumptions and expectations that once legitimized these

practices still prevail at present, and continue to influence the meanings that people attach to contemporary sexual relationships and the expectations that people have in relationships. ... A confluence of old and new is finding expression in many aspects of life, including in the particularities of courtship and sexual relationships. ([Leclerc-Madlala, 2008](#), pp. S22-S23)

Additionally, the equalization of circular migration should equalize the HIV transmission risks that accompany this phenomenon ([Weine and Kashuba, 2012](#)). In contrast, this convergence and the one in men's schooling implies that there is little historical reason to expect differences in economic development along the border. I test both sets of predictions in the next section.

5.2. Data

To compare the effects of the two former institutions on HIV prevalence and economic development today, I use georeferenced individual-level data from up to four waves of the Demographic and Health Surveys (DHS) in Mozambique. Figure 6 shows the reported locations of the survey clusters within 200 km of the institutional border. These locations are slightly displaced to protect respondents' privacy.²²

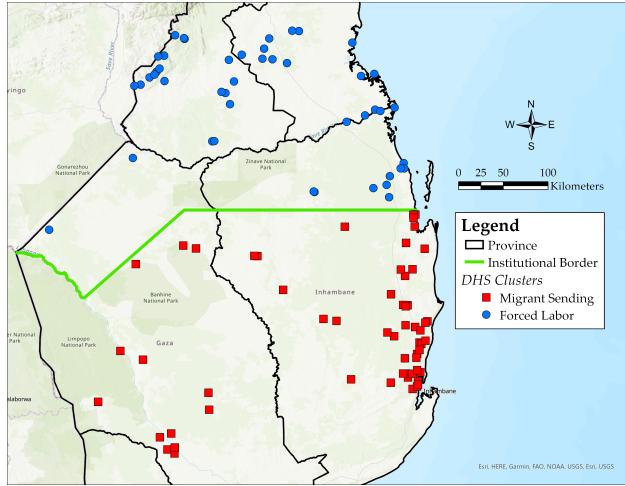
As such, it is possible that four urban clusters along the coast have been displaced into the wrong institution. For this reason and others related to the city's recent history that I discuss in Appendix C, I remove these clusters from the sample. After doing so, all of the remaining ones are in the correct former institution. I discuss in the next section the implications of not knowing their precise locations for the analysis.

The outcome of interest when examining HIV is the result of blood tests for the virus from a random subset of respondents in 2009 and 2015. I restrict this analysis to adults ages 15 to 64.²³ For economic development, the variables of interest are an index of household asset ownership (measured in 2009, 2011, 2015, and 2018), an indicator for

²² Urban clusters are displaced by up to 2 km, 99 percent of rural clusters by up to 5 km, and 1 percent of rural clusters by up to 10 km.

²³ Most studies of HIV prevalence focus on ages 15 to 44 or 49, as this range captures current sexual activity. I use the full adult age range of HIV blood tests in the DHS data because my interest is in the institutions' effects on anyone who was ever sexually active.

Figure 6: Map of Georeferenced DHS Clusters



whether a child is stunted (2011), and the years of schooling an individual has completed (2009, 2011, and 2015).²⁴

5.3. Empirical Strategy

As with the colonial-era analysis, I use an RD design to compare the long-run impact of historical assignment to the migrant-sending institution relative to the forced labor institution. However, the individual-level DHS data and their greater geographic disaggregation allow for several additions to Equation (1). I modify it to be:

$$y_{i,c} = \alpha + \tau \text{MigrantSending}_c + f(\text{Distance}_c) + \text{Lon}_c + \mathbf{X}_i \beta + \delta_t + \epsilon_{i,c} \quad \text{for } c \in B_{\text{MSE}}^* \quad (2)$$

where $y_{i,c}$ is an outcome for individual i in DHS survey cluster c and the first three right-hand side variables are as before. I also include the vector \mathbf{X}_i containing individual-level controls (age, age squared, and a female indicator) and the survey-year fixed effect δ_t . The DHS data have sufficiently many clusters near the border to estimate the [Calonico, Cattaneo and Titiunik \(2014\)](#) MSE-optimal bandwidth, which defines the set of them in B_{MSE}^* . I continue to use a local linear RD specification with a triangular kernel.

²⁴ The index equals a household's quintile in the first principal component of a principal component analysis of its assets (1 = lowest, 5 = highest). A child is considered stunted if its height-for-age z-score using the World Health Organization's Child Growth Standards is less than -2.

5.3.1. Addressing Concerns with Estimation and Inference

An estimation issue arises from the displacement of clusters mentioned earlier. Because the displacement is done randomly, the result is classical measurement error in the running variable. Thus, there is a bias toward zero in the estimated RD coefficients.

For inference, I cluster standard errors by DHS survey cluster. However, two problems can arise from this approach. The first is that the MSE-optimal bandwidths often contain only a “small” number of clusters (i.e., fewer than 30). As a solution, I use the wild cluster bootstrap to calculate p -values as [Cameron, Gelbach and Miller \(2008\)](#) recommend. The second potential concern is spatial autocorrelation, which I again address through [Colella et al. \(2020\)](#) standard errors and [Moran \(1950\)](#) I -statistics.²⁵

6. Effects of the Former Institutions in the Modern Era

Tables [3](#) and [4](#) report the respective RD estimates for HIV prevalence and development outcomes. Figure [7](#) presents RD plots for selected outcomes. They show that today, consistent with the conceptual framework, HIV prevalence is much lower just inside the former migrant-sending institution and there are no differences in development.

6.1. HIV Prevalence

I first examine the spatial distribution of HIV among adults along the institutional border. Table [3](#) Column (1) pools both sexes and shows that adult HIV prevalence drops 10 p.p. just inside the migrant-sending institution. This point estimate is large relative to the 22 percent of the forced labor institution sample who are HIV positive. In addition, the wild cluster bootstrap p -value and the measure of spatial autocorrelation suggest that its statistical significance is not due to false precision.

Splitting the sample by sex in Columns (2) and (3) reveals that this effect is of equal

²⁵ Because there are multiple observations at each site, I collapse individual-level residuals into cluster-level means. I also use a bandwidth of approximately 100 km given the greater density of clusters.

Table 3: HIV Prevalence

	Positive Blood Test		
	Pooled (1)	Women (2)	Men (3)
Migrant Sending	-0.103 (0.049) [0.037]	-0.089 (0.053) [0.041]	-0.097 (0.084) [0.093]
Observations	860	588	212
Clusters	21	22	14
Bandwidth	124.4	128.3	86.5
Wild Cluster Bootstrap p	0.073	0.137	0.458
Spatial Autocorrelation	0.07	-0.05	0.05
Spatial Autocorrelation SD	0.21	0.20	0.19
Forced Labor Mean	0.215	0.214	0.198

Notes: Standard errors clustered by DHS survey cluster are in parentheses. Regressions estimate a local linear RD specification on each side of the border using a triangular weighting kernel and include age, age squared, a female indicator, longitude, and year fixed effects as controls. Specifications use the MSE-optimal bandwidth in kilometers ([Calonico, Cattaneo and Titiunik, 2014](#)).

magnitude for women and men. Figures [7a](#) and [7b](#) provide visual evidence of these sizable differences at the border. However, after accounting for the number of clusters the male estimate is imprecise, likely as a result of the much smaller sample size.

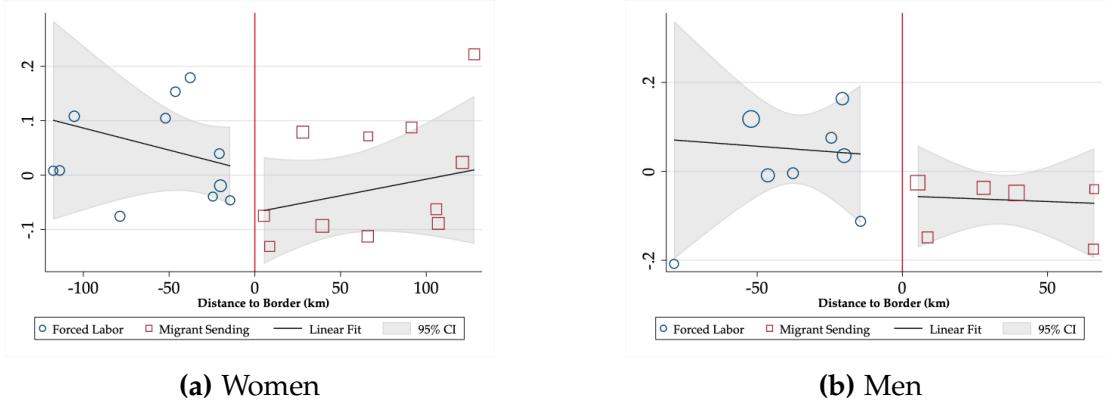
To help rationalize the large effect sizes estimated above, I compare the age profiles of HIV prevalence in the two institutions.²⁶ Specifically, I calculate the mean seroprevalence for each 10-year age group within each sex's MSE-optimal RD bandwidth.²⁷ Figure 8 plots these age profiles. A clear pattern emerges among women: HIV prevalence for every age group is lower (or at a minimum no greater) in the former migrant-sending institution than in the former forced labor region. It is most apparent for women ages 25 to 34, when HIV prevalence peaks in the former forced labor institution. There is a similar pattern among men but the small sample size results in substantial noise.

These age profiles of HIV prevalence are consistent with the [de Oliveira et al. \(2017\)](#) cycle of transmission between and within generations as well as its implications

²⁶ Note also that prevalence is a stock, not a flow, and even small differences in transmission rates can generate large differences in the size of an epidemic ([Viboud, Simonsen and Chowell, 2016](#)).

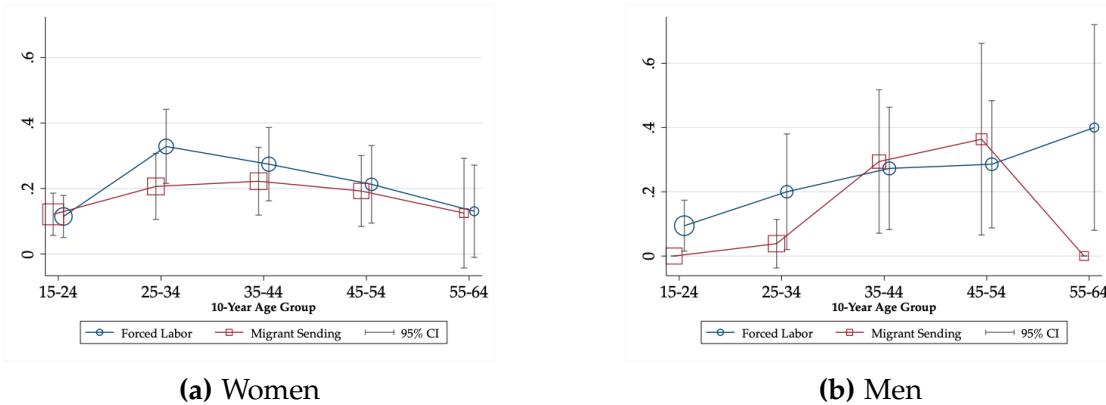
²⁷ The HIV sample is too small to permit reliable RD estimation by sex and age group.

Figure 7: RD Plots for HIV Prevalence



Notes: RD plots show the fraction HIV positive in DHS survey clusters net of age, age squared, longitude, and year fixed effects. The running variable is a cluster's distance to the border. Black lines denote linear trends on each side of the border using a triangular kernel and gray shading indicates 95% confidence intervals. Shape sizes reflect the relative number of adults in a cluster.

Figure 8: Age Profiles of HIV Prevalence



Notes: Plots show the mean HIV prevalence for each sex within a 10-year age group in the two institutions.

Table 4: Economic Development Outcomes

	<i>Assets</i>	<i>Stunting</i>	<i>Years of Schooling</i>	
	Index (1)	Children (2)	Females (3)	Males (4)
Migrant Sending	0.067 (0.322) [0.414]	-0.055 (0.117) [0.132]	0.377 (0.327) [0.281]	0.224 (0.782) [0.795]
Observations	2,513	258	883	815
Clusters	22	15	19	22
Bandwidth	59.6	108.0	64.8	71.5
Wild Cluster Bootstrap p	0.862	0.824	0.302	0.818
Spatial Autocorrelation	-0.19	-0.30	-0.26	-0.16
Spatial Autocorrelation SD	0.15	0.25	0.16	0.15
Forced Labor Mean	3.375	0.376	2.498	3.443

Notes: Standard errors clustered by DHS survey cluster are in parentheses. Regressions estimate a local linear RD specification on each side of the border using a triangular weighting kernel and include age, age squared, a female indicator, longitude, and year fixed effects as controls. Specifications use the MSE-optimal bandwidth in kilometers ([Calonico, Cattaneo and Titiunik, 2014](#)).

discussed in Section 5.1. The peak of women's seroprevalence in the former forced labor institution is both larger than the one in the former migrant-sending region and farther away from the peak for men. Such a pattern could arise from relationships with wider age disparities, as more older men in the highest-prevalence age groups transmit the virus to more younger women. I test for evidence supporting this and other channels in Section 7.

6.2. Economic Development

Next, I compare economic development outcomes in the former institutions. The point estimates in Table 4 are all in the direction of better outcomes in the former migrant-sending region, though for the asset ownership index and male schooling they are a very small percentage of the means in the former forced labor institution. The coefficients for childhood stunting (-6 p.p.) and female schooling (0.38 years) are non-trivial relative to the respective forced labor means (38 percent and 2.5 years) but notably imprecise, especially for the former. Importantly, the RD plots for these outcomes in Appendix A7

suggest that these estimates arise as an artifact of the polynomials' slopes rather than a visible change at the border in their levels.

Taken together, these results fail to provide convincing evidence that economic development changes meaningfully at the border today. They are also consistent with the equalization of circular migration and human capital outcomes in the 1960 census data. The implication is that whatever development differences there were prior to the forced labor institution's abolition have disappeared in the intervening decades.

7. Explaining Differences in HIV Prevalence

In this section, I examine the channels that give rise to lower HIV prevalence in the former migrant-sending institution today. I focus first on age-disparate relationships given the colonial-era marriage market results in Section 4 and the age profiles of seroprevalence in Section 6. I then turn to other risk factors commonly cited in the HIV literature. The main result is that age gaps between relationship and sexual partners are substantially smaller just inside the former migrant-sending institution, and this difference is the main one of note among the potential channels examined.

7.1. Age-Disparate Partnerships

To examine age disparities between marital and sexual partners, I use data from two sources: the IPUMS 10-percent sample of the 2007 Mozambican census and the DHS. The former allows for characteristics of a husband, wife, or live-in partner (henceforth spouse) also in the sample to be attached to an observation, which leads to a dataset with far more linked couples—and likely much more representative ones—than the DHS.²⁸ However, the census lacks information on sexual activity and the data are at a much

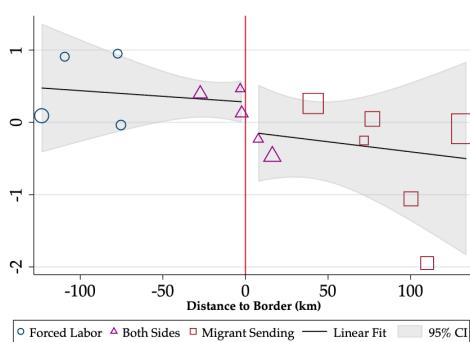
²⁸ Inclusion in the DHS dataset of linked couples requires both partners to be present for and participate in enumeration. Because there are likely important differences between such couples and those with a partner absent from the survey (e.g., they are working outside of the home), selection into the DHS couples sample is a non-trivial consideration.

Table 5: Age-Disparate Partnerships

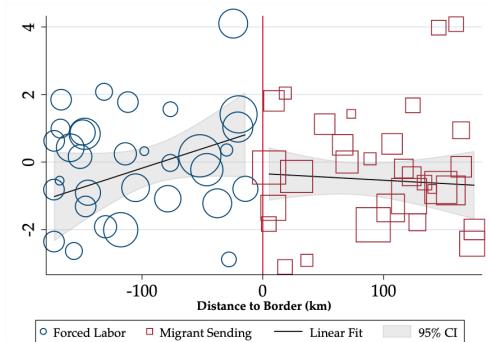
	Census	DHS	
	Spouse	Last Sex Partner	
	Women (1)	Women (2)	Men (3)
Migrant Sending	-0.790 (0.289) [0.220]	-3.129 (1.492) [1.257]	-1.912 (0.828) [0.833]
Observations	9,307	204	300
Clusters	15	14	63
Bandwidth	138.3	56.4	179.0
Wild Cluster Bootstrap p	0.075	0.154	0.081
Spatial Autocorrelation	0.18	0.06	0.12
Spatial Autocorrelation SD	0.21	0.17	0.08
Forced Labor Mean	8.380	7.265	5.110

Notes: Standard errors clustered by DHS survey cluster are in parentheses. Regressions estimate a local linear RD specification on each side of the border using a triangular weighting kernel and include age, age squared, longitude, and year fixed effects as controls. Specifications use the MSE-optimal bandwidth in kilometers ([Calonico, Cattaneo and Titiunik, 2014](#)).

Figure 9: RD Plots for Age Disparities between Partners



(a) Census: Women with Spouse



(b) DHS: Men with Last Sexual Partner

Notes: RD plots show the mean age disparity among the specified group in an administrative post or DHS survey cluster net of age, age squared, longitude, and year fixed effects. The running variable is a cluster's distance to the border. Black lines denote linear trends on each side of the border using a triangular kernel and gray shading indicates 95% confidence intervals. Shape sizes reflect the relative number of adults in a cluster.

coarser geographic resolution (administrative posts, see Appendix A5 for a map).²⁹

The outcome of interest in both datasets is the man's age minus the woman's (henceforth age disparity).³⁰ I examine it between women of any age and their spouses in the census sample to make the closest link possible between a colonial-era and present-day marriage market outcome associated with HIV risk. To connect them directly to sexual behavior, I also study age disparities between individuals of reproductive age (15 to 49) in the DHS data and their most recent sexual partner.³¹

Table 5 reports the results of estimating Equation (2) for age disparities in these datasets and Figure 9 presents RD plots for selected outcomes. Column (1) shows that the age disparity between a woman and her linked spouse in the census is 0.79 years smaller just inside the former migrant-sending institution. This estimate is meaningful relative to the average of 8.38 years in the former forced labor region, and its precision does not seem to stem from the low number of clusters or significant spatial autocorrelation. Figure 9a shows this discontinuity visually.

In Columns (2) and (3) I examine the age disparity in the DHS between reproductive-age individuals and their most recent sexual partners. These estimates for women (-3.13 years) and men (-1.91 years) are even larger relative to the forced labor means (7.27 years and 5.11 years), though accounting for the number of clusters slightly reduces the precision of the women's estimate. Figure 9b shows the RD plot for men.

7.2. Risk Factors Associated with Age-Disparate Partnerships

While age-disparate partnerships can be HIV risk factors in themselves, they are also associated with other behaviors facilitating transmission of the virus in Sub-Saharan

²⁹ As in the 1960 data, these administrative boundaries do not align with the former institutional border. The effects on the estimation are the same as I discussed in Section 3.2.1, and in the RD plots I take the same approach to the administrative posts with area on both sides of the border.

³⁰ I structure the data so that each observation is a woman linked with her spouse to account for polygyny. Thus, each woman only appears once but a man can be linked to multiple women.

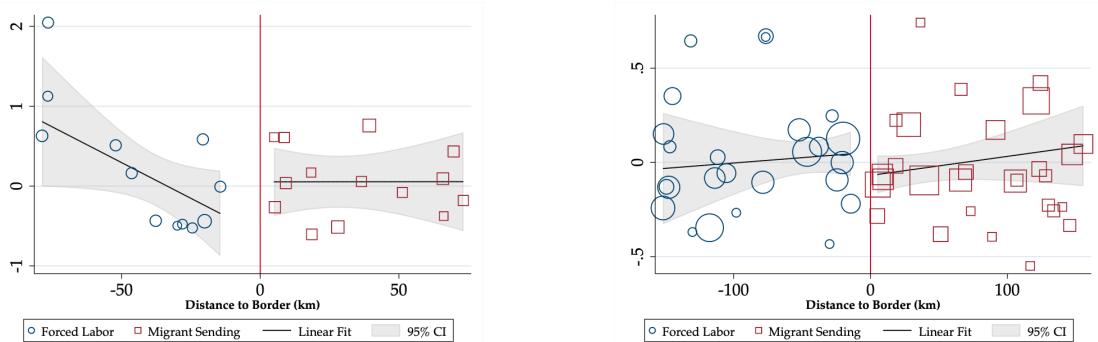
³¹ Ninety-six percent of women and 89 percent of men in the DHS reported their most recent sexual partner was a spouse or boyfriend/girlfriend, implying respondents should know this person's age.

Table 6: HIV Risk Factors Associated with Age Disparities

	<i>Age at First Sex</i>	<i>Has Partners Concurrently</i>	<i>Condom Used Last Sex</i>	
	Women (1)	Men (2)	Women (3)	Men (4)
Migrant Sending	0.813 (0.379) [0.403]	-0.157 (0.086) [0.081]	0.001 (0.045) [0.039]	0.065 (0.058) [0.061]
Observations	603	250	375	136
Clusters	26	54	28	26
Bandwidth	86.8	156.1	96.1	85.3
Wild Cluster Bootstrap p	0.169	0.153	0.986	0.448
Spatial Autocorrelation	0.01	0.07	-0.28	0.14
Spatial Autocorrelation SD	0.15	0.09	0.16	0.15
Forced Labor Mean	16.13	0.258	0.056	0.069

Notes: Standard errors clustered by DHS survey cluster are in parentheses. Regressions estimate a local linear RD specification on each side of the border using a triangular weighting kernel and include age, age squared, longitude, and year fixed effects as controls. Specifications use the MSE-optimal bandwidth in kilometers ([Calonico, Cattaneo and Titiunik, 2014](#)).

Figure 10: RD Plots for HIV Risk Factors Associated with Age Disparities



(a) Women: Age at First Sex

(b) Men: Has Partners Concurrently

Notes: RD plots show mean of the specified outcome in DHS survey clusters net of age, age squared, longitude, and year fixed effects. The running variable is a cluster's distance to the border. Black lines denote linear trends on each side of the border using a triangular kernel and gray shading indicates 95% confidence intervals. Shape sizes reflect the relative number of adults in a cluster.

Africa. These risk factors include the woman having her first sexual intercourse at a younger age, having a male partner who is in a concurrent relationship, and not using a condom ([Evans et al., 2019](#); [Mabaso et al., 2021](#); [Schaefer et al., 2017](#)).

I test for differences in these outcomes along the border in Table 6 and present RD plots for selected ones in Figure 10. The estimate in Column (1) reflects a woman's age of sexual debut being 0.81 years later in the former migrant-sending region compared to a mean in the former forced labor area of 16.1 years. In addition, Column (2) shows that men in the former migrant-sending region are 16 p.p. less likely to have concurrent sexual partners, which is substantial compared to the forced labor mean of 26 percent. Although both of these estimates are slightly imprecise, the changes at the border are visually apparent in Figure 10, especially for concurrent partners.

In contrast, the evidence on condom use in the most recent sexual intercourse is less compelling. In Column (3), there is a null RD estimate for women relative to a 5.6-p.p. mean in the forced labor region. The coefficient for men in Column (4) is much larger—6.5 p.p. relative to a forced labor mean of 7 percent—but very imprecise.

7.3. Other HIV Risk Factors

Finally, I turn to other important risk factors in Sub-Saharan Africa to determine whether partner age disparities and associated behaviors are the main channels for the HIV effect in Section 6. Drawing from the literature on the virus' spread across the continent, I create indicator variables for: having a genital ulcer in the past 12 months, a polygynous partnership, a woman having ever been forced to engage in sexual activity, a man having ever paid for sex, a woman being able to decide alone on her own healthcare, and a man having been medically circumcised.

Table 7 shows the results of estimating Equation (2) for these outcomes. The only ones with effects pointing in the direction of lower HIV prevalence in the former migrant-sending region are fewer women being forced to engage in sex (-2.4 p.p. relative to a

Table 7: Ruling Out Other HIV Risk Factors

	<i>Genital Ulcer in Last Year</i>		<i>Polygynous Partnership</i>		<i>Forced Sex</i>	<i>Paid for Sex</i>	<i>Health Decider</i>	<i>Medical Circumc.</i>
	Women (1)	Men (2)	Women (3)	Men (4)	Women (5)	Men (6)	Women (7)	Men (8)
Migrant Sending	0.061 (0.028) [0.030]	0.027 (0.014) [0.017]	-0.012 (0.045) [0.054]	0.017 (0.049) [0.057]	-0.024 (0.043) [0.030]	0.015 (0.081) [0.062]	-0.057 (0.094) [0.121]	0.031 (0.175) [0.079]
Observations	414	235	466	202	220	196	332	443
Clusters	19	26	29	42	21	38	22	54
Bandwidth	54.7	85.0	98.2	130.8	90.7	148.2	92.8	152.8
Wild Cluster Bootstrap <i>p</i>	0.077	0.216	0.867	0.733	0.630	0.838	0.650	0.846
Spatial Autocorrelation	-0.17	-0.13	0.07	-0.38	-0.25	0.21	0.06	-0.15
Spatial Autocorrelation SD	0.16	0.15	0.15	0.11	0.16	0.11	0.18	0.09
Forced Labor Mean	0.005	0.008	0.090	0.050	0.087	0.085	0.266	0.192

Notes: Standard errors clustered by DHS survey cluster are in parentheses. Regressions estimate a local linear RD specification on each side of the border using a triangular weighting kernel and include age, age squared, longitude, and year fixed effects as controls. Specifications use the MSE-optimal bandwidth in kilometers ([Calonico, Cattaneo and Titiunik, 2014](#)).

forced labor mean of 8.7 percent) and more men who are medically circumcised (3.1 p.p. relative to a forced labor mean of 19.2 percent), but both estimates are highly imprecise. Additionally, the largest and only precisely estimated effect—6.1 p.p. more women having a genital ulcer in the last year relative to a forced labor mean of 0.5 percent—*increases* HIV risk in the former migrant-sending region. Taken together, the results suggest that these commonly-cited risk factors do not explain the drop in HIV prevalence just inside the former migrant-sending area.

8. Conclusion

The results in this paper show that two institutions common throughout colonial Africa can have markedly different impacts on HIV prevalence today, likely through their lasting effects on dating and marriage markets. However, there is little colonial-era or present-day evidence for lasting effects on economic development.

References

- Acemoglu, Daron, and James A. Robinson.** 2012. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. New York:Crown Business. [1]
- Acemoglu, Daron, Simon Johnson, and James A. Robinson.** 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review*, 91(5): 1369–1401. [4]
- Alexopoulou, Kleoniki, and Dácil Juif.** 2017. "Colonial State Formation without Integration: Tax Capacity and Labour Regimes in Portuguese Mozambique (1890s-1970s)." *International Review of Social History*, 62(2): 215–252. [2]
- Allina, Eric.** 2012. *Slavery by Any Other Name: African Life under Company Rule in Colonial Mozambique*. Charlottesville, VA:University of Virginia Press. [9, 10, 11]
- Alsan, Marcella, and Marianne Wanamaker.** 2018. "Tuskegee and the Health of Black Men." *Quarterly Journal of Economics*, 133(1): 407–455. [5]
- Amin, Samir.** 1972. "Underdevelopment and Dependence in Black Africa: Origins and Contemporary Forms." *Journal of Modern African Studies*, 10(4): 503–524. [1, 2]
- Anderson, Siwan.** 2018. "Legal Origins and Female HIV." *American Economic Review*, 108(6): 1407–1439. [5]
- Audet, Carolyn M., Janeen Burlison, Troy D. Moon, Mohsin Sidat, Alfredo E. Vergara, and Sten H. Vermund.** 2010. "Sociocultural and Epidemiological Aspects of HIV/AIDS in Mozambique." *BMC International Health and Human Rights*, 10: 15. [12]
- Banerjee, Abhijit, and Lakshmi Iyer.** 2005. "History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India." *American Economic Review*, 95(4): 1190–1213. [4]
- Barakat, Sultan, Margaret Chard, Tim Jacoby, and William Lume.** 2002. "The Composite Approach: Research Design in the Context of War and Armed Conflict." *Third World Quarterly*, 23(5): 991–1003. [12]
- Bertocchi, Graziella, and Arcangelo Dimico.** 2019. "The Long-Term Determinants of Female HIV Infection in Africa: The Slave Trade, Polygyny, and Sexual Behavior." *Journal of Development Economics*, 140: 90–105. [5]
- Cagé, Julia, and Valeria Rueda.** 2020. "Sex and the Mission: The Conflicting Effects of Early Christian Missions on HIV in sub-Saharan Africa." *Journal of Demographic Economics*, 86(3): 213–257. [5]
- Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik.** 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica*, 82(6): 2295–2326. [14, 17, 25, 27, 29, 31, 33, 35]

- Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller.** 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *Review of Economics and Statistics*, 90(3): 414–427. [16, 26]
- Cattaneo, Matias D., Nicolás Idrobo, and Rocío Titiunik.** 2019. *A Practical Introduction to Regression Discontinuity Designs: Foundations*. Cambridge, UK:Cambridge University Press. [15]
- Chiappori, Pierre-André, Murat Iyigun, and Yoram Weiss.** 2009. "Investment in Schooling and the Marriage Market." *American Economic Review*, 99(5): 1689–1713. [5]
- Clarence-Smith, Gervase.** 1985. *The Third Portuguese Empire, 1825-1975: A Study in Economic Imperialism*. Manchester:Manchester University Press. [8]
- Colella, Fabrizio, Rafael Lalive, Seyhun Orcan Sakalli, and Mathias Thoenig.** 2020. "Inference with Arbitrary Clustering." [15, 19, 26]
- Crush, Jonathan, Alan Jeeves, and David Yudelman.** 1991. *South Africa's Labor Empire: A History of Black Migrancy to the Gold Mines*. Boulder, CO:Westview Press. [42]
- Darch, Colin.** 1983. "Notas sobre Fontes Estatísticas Oficiais Referente à Economia Colonial Moçambicana: Uma Crítica Geral." *Estudos Moçambicanos*, 4: 103–125. [12]
- Dell, Melissa.** 2010. "The Persistent Effects of Peru's Mining Mita." *Econometrica*, 78(6): 1863–1903. [4]
- Dell, Melissa, and Benjamin A. Olken.** 2020. "The Development Effects of the Extractive Colonial Economy: The Dutch Cultivation System in Java." *Review of Economic Studies*, 87(1): 164–203. [4]
- de Oliveira, Tulio, Ayesha B. M. Kharsany, Tiago Gräf, Cherie Cawood, David Khanyile, Anneke Grobler, Adrian Puren, Savathree Madurai, Cheryl Baxter, Quarraisha Abdool Karim, and Salim S. Abdool Karim.** 2017. "Transmission Networks and Risk of HIV Infection in KwaZulu-Natal, South Africa: A Community-Wide Phylogenetic Study." *The Lancet HIV*, 4(1): E41–E50. [3, 23, 27]
- Dinkelmann, Taryn, and Martine Mariotti.** 2016. "The Long-Run Effects of Labor Migration on Human Capital Formation in Communities of Origin." *American Economic Journal: Applied Economics*, 8(4): 1–35. [3, 4]
- Direcção Provincial dos Serviços de Estatística.** 1966. *III Recenseamento Geral da População na Província de Moçambique*. Lourenço Marques:Instituto Nacional de Estatística. [12, 45]
- Dokkum, André van.** 2020. *Nationalism and Territoriality in Barue and Mozambique: Independence, Belonging, Contradiction*. Leiden, The Netherlands:Brill. [6]

Dwyer-Lindgren, Laura, Michael A. Cork, Amber Sligar, Krista M. Steuben, Kate F. Wilson, Naomi R. Provost, Benjamin K. Mayala, John D. VanderHeide, Michael L. Collison, Jason B. Hall, and et al. 2019. "Mapping HIV Prevalence in Sub-Saharan Africa between 2000 and 2017." *Nature*, 570: 189–193. [5]

Evans, Meredith, Kathryn Risher, Nompumelelo Zungu, Olive Shisana, Sizulu Moyo, David C. Celentano, Brendan Maughan-Brown, and Thomas M. Rehle. 2019. "Age-Disparate Sex and HIV Risk for Young Women from 2002 to 2012 in South Africa." *Journal of the International AIDS Society*, 19(1): 21310. [4, 34]

Fernández Cebrián, Pablo. 2021. "Withdrawal of the State: The Provision of Primary Schooling in Mozambique under the Indigenato." [11, 22]

Frimmel, Hartwig E. 2019. "The Witwatersrand Basin and Its Gold Deposits." In *The Archean Geology of the Kaapvaal Craton, Southern Africa. Regional Geology Reviews*, , ed. Alfred Kröner and Axel Hoffman. Cham, CHE:Springer. [8]

Fuller, Charles Edward. 1955. "An Ethnohistoric Study of Continuity and Change in Gwambe Culture." PhD diss. Northwestern University, Evanston, IL. [10]

Gelman, Andrew, and Guido Imbens. 2019. "Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs." *Journal of Business & Economic Statistics*, 37(3): 447–456. [15]

Gengenbach, Heidi. 2010. *Binding Memories: Women as Makers and Tellers of History in Magude, Mozambique*. New York:Columbia University Press. [44]

Govêrno Geral de Moçambique. 1942. *Moçambique: Documentário Trimestral*. Vol. 31, Lourenço Marques:Imprensa Nacional de Moçambique. [43]

Great Britain Foreign Office. 1901. "Portuguese Decree, Extending the Concessions Granted to the Mozambique Company - Lisbon, December 22, 1893." In *British and Foreign State Papers, 1896-1897*. Vol. 89, , ed. Augustus H. Oakes and Willoughby Maycock, 601–602. London:H.M.S.O. [6]

Griffith, Daniel A., and Giuseppe Arbia. 2010. "Detecting Negative Spatial Autocorrelation in Georeferenced Random Variables." *International Journal of Geographical Information Science*, 24(3): 417–437. [16]

Guthrie, Zachary Kagan. 2016. "Forced Volunteers: The Complexities of Coercion in Central Mozambique, 1941-1961." *International Journal of African Historical Studies*, 49(2): 195–212. [11]

Guthrie, Zachary Kagan. 2018. *Bound for Work: Labor, Mobility, and Colonial Rule in Central Mozambique, 1940-1965*. Charlottesville, VA:University of Virginia Press. [9, 10]

Harries, Patrick. 1994. *Work, Culture, and Identity: Migrant Laborers in Mozambique and South Africa, c. 1860-1910*. Portsmouth, NH:Heinemann. [8, 10]

- Harris, Marvin.** 1959. "Labour Emigration among the Moçambique Thonga: Cultural and Political Factors." *Africa*, 29(1): 50–66. [7]
- Harrison, Graham.** 1998. "Quantos Somos? The Second National Population Census of Mozambique." *Review of African Political Economy*, 25(75): 124–132. [12]
- Havik, Philip J.** 2013. "Colonial Administration, Public Accounts and Fiscal Extraction: Policies and Revenues in Portuguese Africa (1900-1960)." *African Economic History*, 41: 159–221. [12]
- Helgesson, Alf.** 1994. *Church, State and People in Mozambique: An Historical Study with Special Emphasis on Methodist Developments in the Inhambane Region*. Uppsala:Swedish Institute of Missionary Research. [11]
- Henrich, Joseph, Robert Boyd, and Peter J. Richerson.** 2012. "The Puzzle of Monogamous Marriage." *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1589): 657–669. [4]
- Iliffe, John.** 2006. *The African AIDS epidemic: A History*. Athens, OH:Ohio University Press. [12]
- Irons, William.** 1983. "Human Female Reproductive Strategies." In *Social Behavior of Female Vertebrates*. , ed. Samuel Wasser, 169–213. New York:Academic Press. [4]
- Isaacman, Allen, Michael Stephen, Yusuf Adam, Maria João Homen, Eugenio Macamo, and Augustinho Pililão.** 1980. "'Cotton Is the Mother of Poverty": Peasant Resistance to Forced Cotton Production in Mozambique, 1938-1961." *International Journal of African Historical Studies*, 13(4): 581–615. [11]
- Junod, Henri A.** 1912. *The Life of a South African Tribe*. Vol. 1, London:David Nutt. [7, 10]
- Kelly, Morgan.** 2020. "Understanding Persistence." [14, 15]
- Khanna, Gaurav, Caroline Theoharides, and Dean Yang.** 2020. "Abundance from Abroad: Migrant Income and Long-Run Economic Development." [5]
- Leclerc-Madlala, Suzanne.** 2008. "Age-Disparate and Intergenerational Sex in Southern Africa: The Dynamics of Hypervulnerability." *AIDS*, 22(suppl. 4): S17–S25. [4, 24]
- Lowes, Sara, and Eduardo Montero.** 2021a. "Concessions, Violence, and Indirect Rule: Evidence from the Congo Free State." [4]
- Lowes, Sara, and Eduardo Montero.** 2021b. "The Legacy of Colonial Medicine in Central Africa." *American Economic Review*, 111(4): 1284–1314. [5]
- Mabaso, Musawenkosi, Lungelo Mlangeni, Lehlogonolo Makola, Olanrewaju Oladimeji, Inbarani Naidoo, Yogandra Naidoo, Buyisile Chibi, Khangelani Zuma, and Leickness Simbayi.** 2021. "Factors Associated with Age-Disparate Sexual Partnerships among Males and Females in South Africa: A Multinomial Analysis

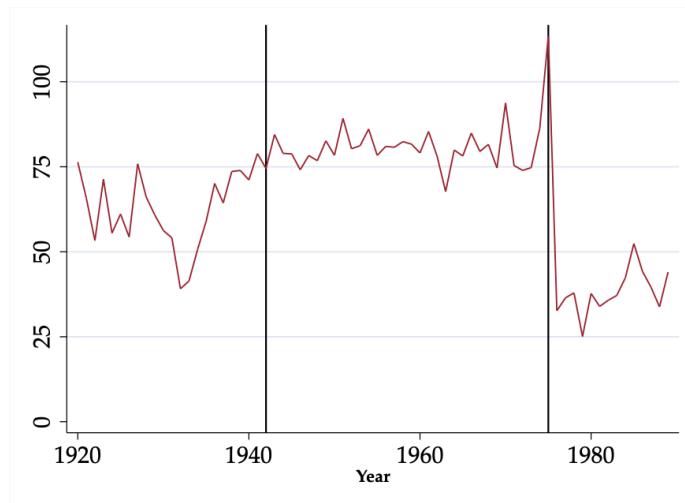
- of the 2012 National Population-Based Household Survey Data." *Emerging Themes in Epidemiology*, 18: 3. [4, 34]
- Michalopoulos, Stelios, and Elias Papaioannou.** 2014. "National Institutions and Subnational Development." *Quarterly Journal of Economics*, 129(1): 151–213. [4]
- Michalopoulos, Stelios, and Elias Papaioannou.** 2016. "The Long-Run Effects of the Scramble for Africa." *American Economic Review*, 106(7): 1802–1848. [4]
- Ministério do Ultramar.** 1959. "Carta de Moçambique." [13]
- Moran, P. A. P.** 1950. "Notes on Continuous Stochastic Phenomena." *Biometrika*, 37(1-2): 17–23. [16, 26]
- Morier-Genoud, Éric.** 2019. *Catholicism and the Making of Politics in Central Mozambique, 1940-1986*. Rochester:University of Rochester Press. [11]
- Mozambique News Agency.** 1999. "Huge Growth at Vilankulo Airport." *AIM Reports*. [49]
- Mozambique News Agency.** 2000. "Tourism Brings in \$32 Million." *AIM Reports*. [49]
- Mozambique News Agency.** 2011. "President Inaugurates New Terminal at Vilankulo Airport." *AIM Reports*. [49]
- Murdock, George Peter.** 1959. *Africa: Its Peoples and Their Culture History*. New York:McGraw-Hill Book Company. [16, 17]
- Méndez-Chacón, Esteban, and Diana Van Patten.** 2021. "Multinationals, Monopsony, and Local Development: Evidence from the United Fruit Company." [4]
- Newitt, Malyn.** 1995. *A History of Mozambique*. Bloomington, IN:University of Indiana Press. [5, 6, 7, 8, 11]
- Newitt, M. D. D.** 1969. "The Portuguese on the Zambezi: An Historical Interpretation of the Prazo System." *Journal of African History*, 10(1): 67–85. [6]
- Portugal.** 1900. *Collecção Official de Legislação Portugueza: Anno de 1899*. Lisboa:Imprensa Nacional. [8]
- Repartição Nacional de Estatística.** 1942. *Censo da População em 1940*. Vol. 2-5, Lourenço Marques:Imprensa Nacional de Moçambique. [12, 45]
- Rita-Ferreira, A.** 1960. "Labour Emigration among the Moçambique Thonga: Comments on a Study by Marvin Harris." *Africa*, 30(2): 141–152. [7]
- Roberts, Richard.** 2017. "Coerced Labor in Twentieth-Century Africa." In *The Cambridge World History of Slavery*. , ed. David Eltis, Stanley L. Engerman, Seymour Drescher and David Richardson, 583–609. Cambridge:Cambridge University Press. [1]

- Ross, Edward Alsworth.** 1925. *Report on Employment of Native Labor in Portuguese Africa*. New York:The Abbott Press. [10]
- Saldanha, F.** 1940. "Carta da Colónia de Moçambique: Segundo a Sua Actual Divisão Administrativa." [13]
- Schaefer, Robin, Simon Gregson, Jeffrey W. Eaton, Owen Mugurungi, Rebecca Rhead, Albert Takaruza, Rufurwokuda Maswera, and Constance Nyamukapa.** 2017. "Age-Disparate Relationships and HIV Incidence in Adolescent Girls and Young Women: Evidence from Zimbabwe." *AIDS*, 31(10): 1461–1470. [3, 23, 34]
- Smith, Alan K., and Gervase Clarence Smith.** 1985. "Angola and Mozambique, 1870–1905." In *The Cambridge History of Africa*. Vol. 6, , ed. Roland Oliver and G. N. Sanderson, 493–521. Cambridge:Cambridge University Press. [5]
- Stichter, Sharon.** 1985. *Migrant Laborers*. Cambridge, UK:Cambridge University Press. [5]
- Transvaal Chamber of Mines.** 1946. "Map of Southern Africa Showing the Recruiting Systems of the Native Recruiting Corporation and the Witwatersrand Native Labour Association." [7]
- Vail, Leroy.** 1976. "Mozambique's Chartered Companies: The Rule of the Feeble." *Journal of African History*, 17(3): 389–416. [6]
- Viboud, Cécile, Lone Simonsen, and Geraldo Chowell.** 2016. "A Generalized-Growth Model to Characterize the Early Ascending Phase of Infectious Disease." *Epidemics*, 15: 27–37. [27]
- Weine, Stevan M., and Adrianna B. Kashuba.** 2012. "Labor Migration and HIV Risk: A Systematic Review of the Literature." *AIDS and Behavior*, 16(6): 1605–1621. [24]
- Weinstein, Jeremy M.** 2006. *Inside Rebellion: The Politics of Insurgent Violence*. New York:Cambridge University Press. [12]
- Wilson, Francis.** 1972. *Labour in the South African Gold Mines, 1911-1969*. New York:Cambridge University Press. [8]
- Yang, Dean.** 2008. "International Migration, Remittances and Household Investment: Evidence from Philippine Migrants' Exchange Rate Shocks." *Economic Journal*, 118(528): 591–630. [3]

Appendix A. Additional Figures

A1. Annual Numbers of Witwatersrand Mine Workers from Southern Mozambique

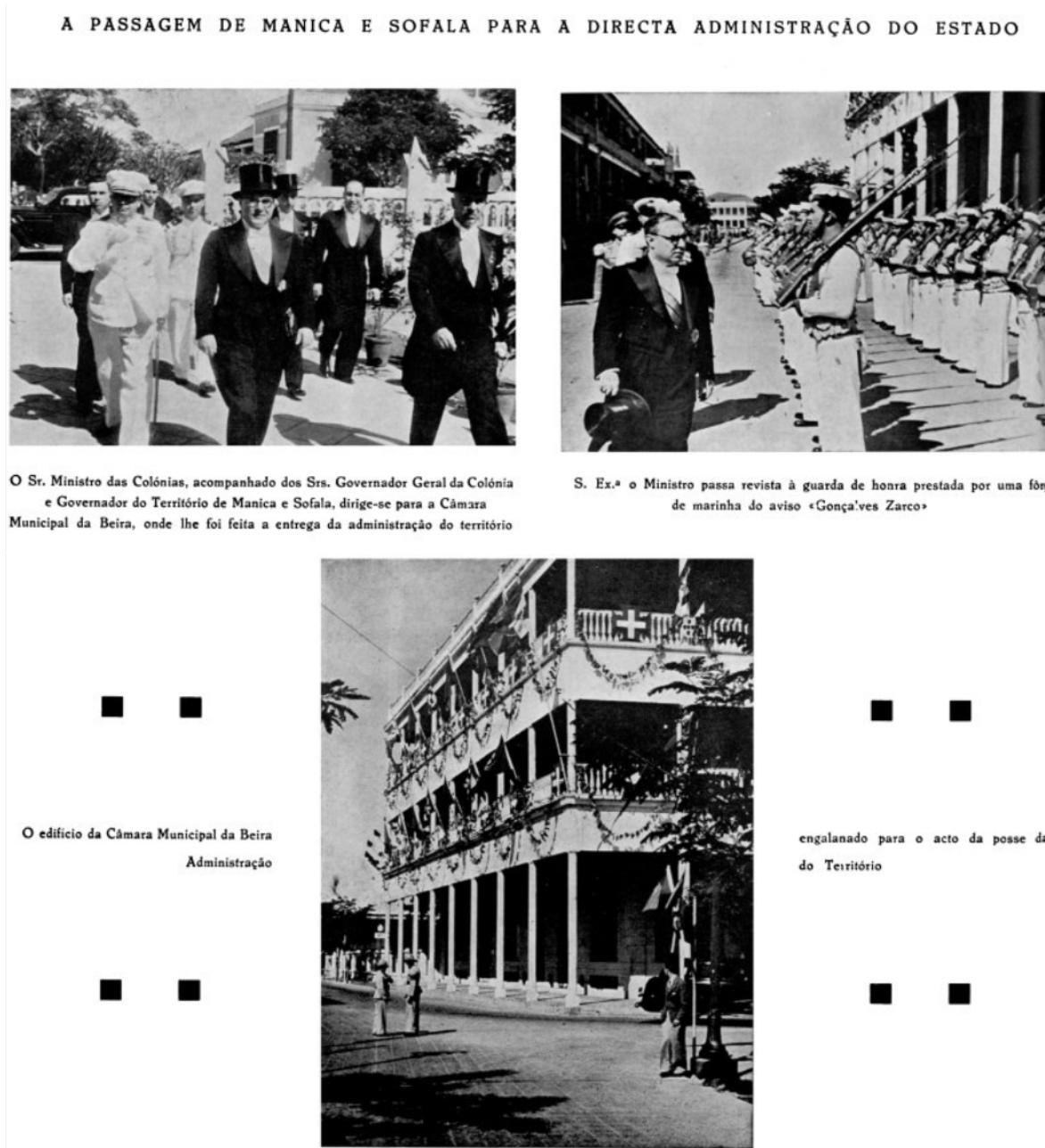
Figure A1: Southern Mozambican Men Received by Mines (000s), 1920-89 [9, 11]



Notes: Data are from the annual reports of the Witwatersrand Native Labour Associated (as cited in [Crush, Jeeves and Yudelman, 1991](#)). The black line in 1942 denotes the end of the forced labor institution, and the black line in 1975 denotes Mozambique's independence from Portugal and deterioration of relations with South Africa.

A2. Handover of the Forced Labor Region to the Colonial State, July 1942

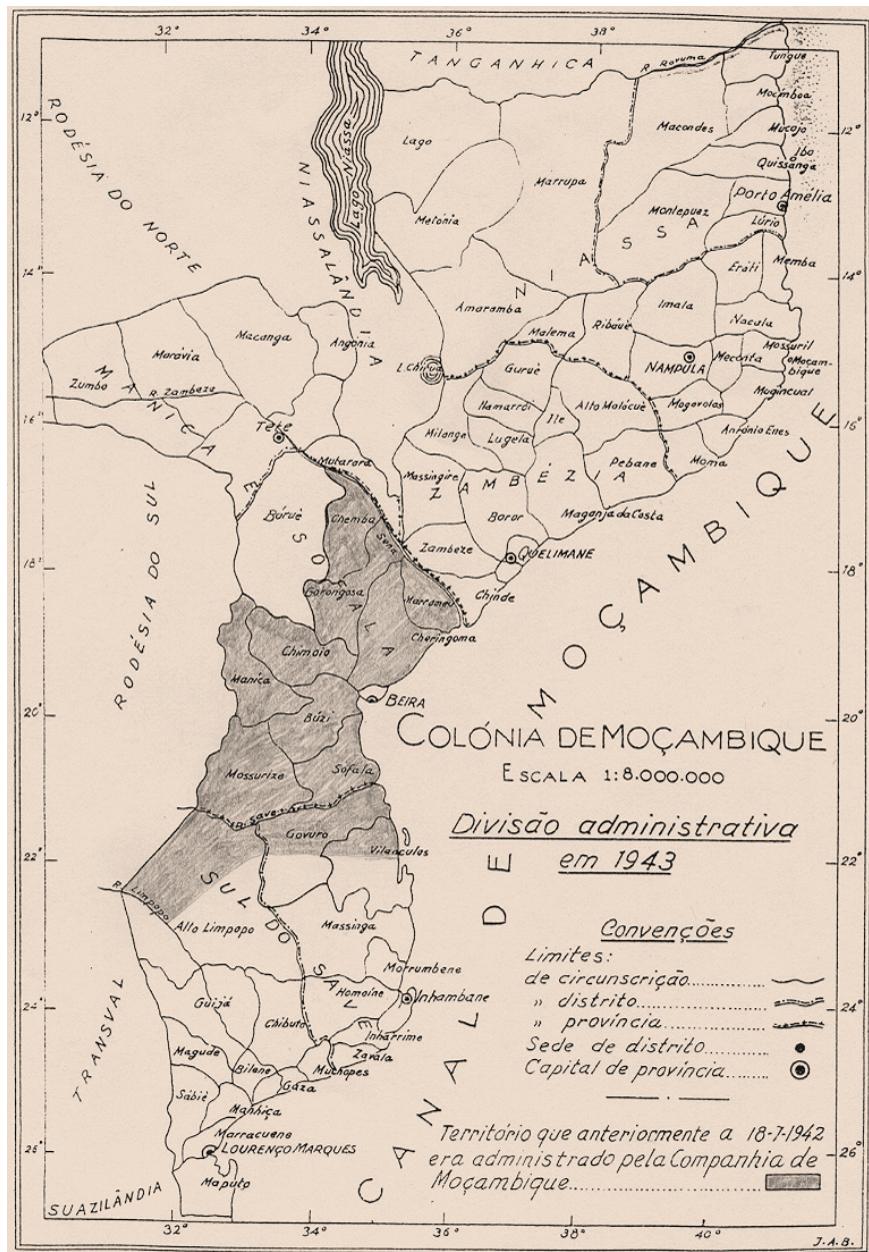
Figure A2: "The Passing of Manica e Sofala to Direct Administration by the State" [11]



Notes: Images are taken from [Govêrno Geral de Moçambique \(1942, p. 108\)](#), courtesy of the *Portal das Memórias de África e do Oriente*. The top left image shows the Minister of Colonies, Governor General of Mozambique, and Governor of the Territory of Manica e Sofala (the forced labor region) walking to the handover ceremony. The top right image shows the Minister reviewing an honor guard at the ceremony. The bottom image shows the Beira City Council building decorated for "the act of taking possession of the [forced labor] territory."

A3. Map of Mozambique after the End of the Forced Labor Institution

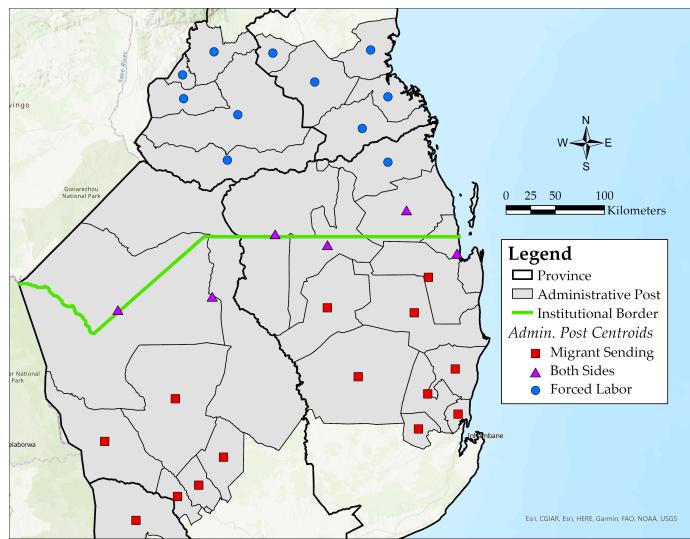
Figure A3: Administrative Divisions of Mozambique, 1943 [11]



Notes: Map is taken from Gengenbach (2010). The shaded area shows the former forced labor region overlaid onto the redrawn administrative boundaries.

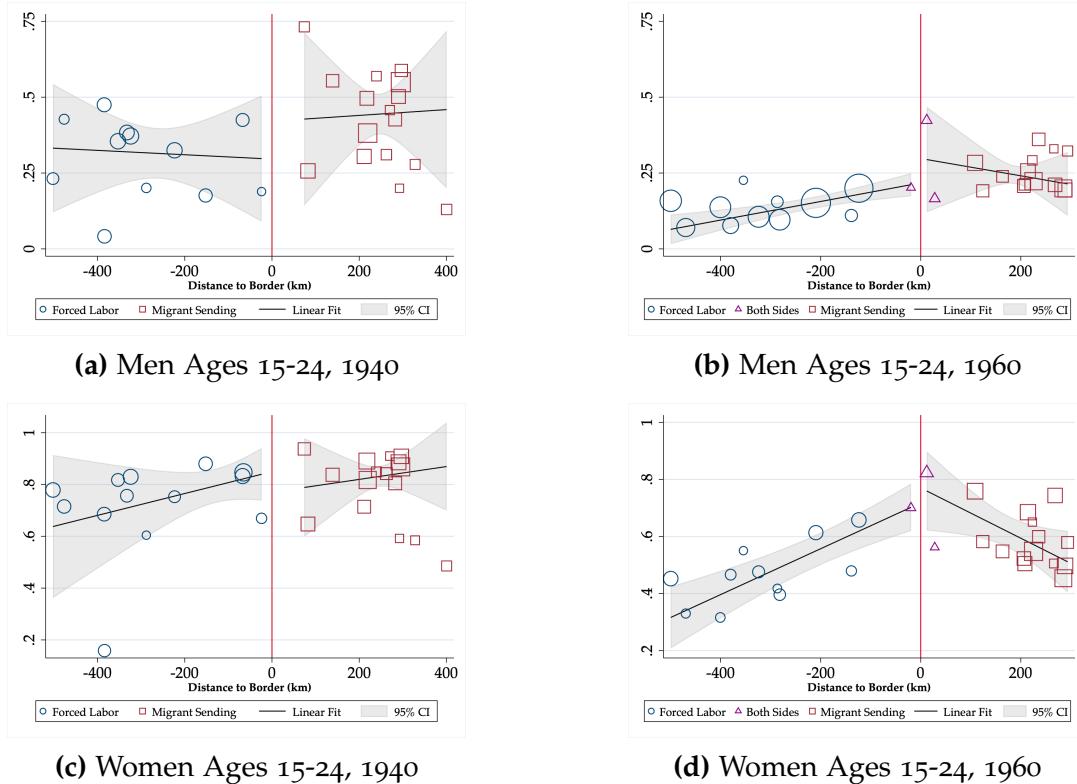
A5. Map of Administrative Posts

Figure A5: Map of Administrative Posts [30]



A6. Additional Colonial-Era RD Plots

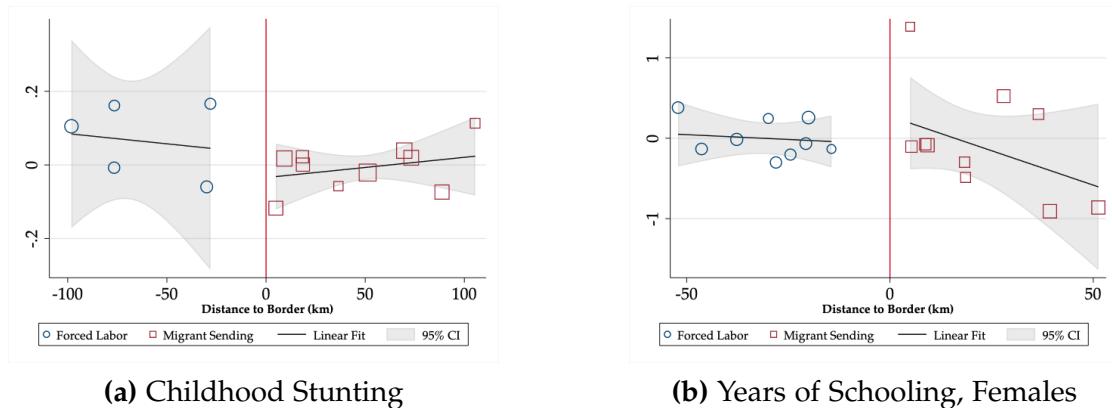
Figure A6: Marriage Rates by Sex for Young Adults [21]



Notes: RD plots show the share in a group among the relevant population in a district. Local linear trends and 95% confidence intervals are estimated on each side of the institutional border using a triangular kernel and weighting by the relevant population. The running variable is distance in kilometers to the border. Data sources, variable definitions, and RD bandwidth selection criteria are in Section 3.1.

A7. Additional DHS RD Plots

Figure A7: Additional RD Plots for Economic Development [29]



Notes: RD plots show mean of the specified outcome in DHS survey clusters net of age, age squared, longitude, and year fixed effects. The running variable is a cluster's distance to the border. Black lines denote linear trends on each side of the border using a triangular kernel and gray shading indicates 95% confidence intervals. Shape sizes reflect the relative number of adults in a cluster.

Appendix B. Additional Tables

B1. Additional Colonial-Era Results

Table B1: Marriage Rates by Sex and Age Group [21]

Panel A: 2 Years before Forced Labor Institution Abolished (1940)						
	Men Married			Women Married		
	Ages 15-24 (1)	Ages 25-34 (2)	Ages 35-44 (3)	Ages 15-24 (4)	Ages 25-34 (5)	Ages 35-44 (6)
Migrant Sending	0.227 (0.145)	0.210 (0.062)	0.024 (0.049)	-0.015 (0.115)	0.024 (0.024)	-0.012 (0.017)
Observations	28	28	28	28	28	28
Bandwidth (km)	[-503, 401]	[-503, 401]	[-503, 401]	[-503, 401]	[-503, 401]	[-503, 401]
Forced Labor Mean	0.314	0.835	0.938	0.727	0.950	0.959

Panel B: 18 Years after Forced Labor Institution Abolished (1960)						
	Men Married			Women Married		
	Ages 15-24 (1)	Ages 25-34 (2)	Ages 35-44 (3)	Ages 15-24 (4)	Ages 25-34 (5)	Ages 35-44 (6)
Migrant Sending	0.052 (0.064)	0.130 (0.066)	0.055 (0.028)	0.038 (0.072)	0.103 (0.051)	0.031 (0.014)
Observations	27	27	27	27	27	27
Bandwidth (km)	[-500, 294]	[-500, 294]	[-500, 294]	[-500, 294]	[-500, 294]	[-500, 294]
Forced Labor Mean	0.143	0.664	0.909	0.506	0.900	0.965

Notes: RD plots show the outcomes in each district. Local linear trends and 95% confidence intervals are estimated on each side of the institutional border using a triangular kernel and weighting by the relevant population. The running variable is distance in kilometers to the border. Plots for a variable have the same y-axis scale across years to facilitate comparisons. Data sources, variable definitions, and RD bandwidth selection criteria are in Section 3.1.

Appendix C. Excluding DHS Clusters within 2 km of the Border [24]

These clusters are in the beachside resort city of Vilankulo, which has been popular with international tourists since the 1990s and has had multimillion-dollar infrastructure upgrades in the last decade (Mozambique News Agency, 1999, 2000, 2011). While this recent history is unrelated to the differences between the institutions, it may have affected outcomes of interest for clusters immediately on one side of the border.