

# **Colonial Institutions, Marriage Markets, and HIV: Evidence from Mozambique<sup>\*</sup>**

Jon Denton-Schneider<sup>†</sup>

*University of Michigan*

October 15, 2021

[Most Recent Version Here](#)

## **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 (1897-1975) and another that heavily restricted their mobility (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 mobility-restricting 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, Eduardo Montero, Ach Adhvaryu, and Dean Yang for their thoughtful guidance and encouragement. I thank them and James Allen, George Alter, Omar Galarraga, Elisa Maffioli, Jennifer Mayo, Alexander Persaud, Ana Reynoso, Elyce Rotella, and conference and seminar participants at Michigan, the NBER Summer Institute 2021 (Development of the American Economy), the 2019 Economic History Association Annual Meeting (Atlanta), and DemSemX for helpful comments and suggestions. 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.

<sup>†</sup>University of Michigan, Department of Economics, 238 Lorch Hall, 611 Tappan Ave, Ann Arbor, MI 48109. Email: [jdds@umich.edu](mailto:jdds@umich.edu). Website: [jondentonschneider.com](http://jondentonschneider.com).

## 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 pushed an "army of short-term male [migrant] labor" to work in mines and on settler farms, while in the Congo Basin it was the *concession*, or a grant of "land (and the Africans living on [it]) to private companies" that heavily restricted the population's mobility ([Roberts, 2017](#), p. 585).<sup>1</sup> Figure 1 shows the present-day countries in which these two were the primary regimes.

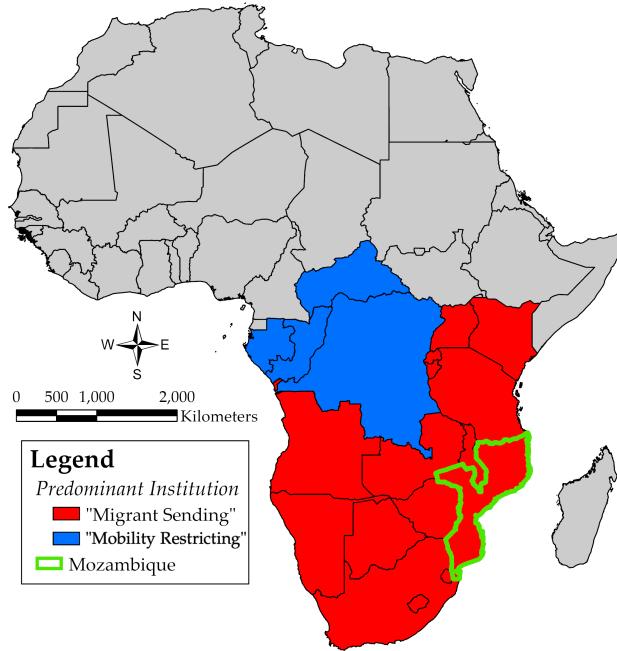
Because these migrant-sending and mobility-restricting institutions organized much of the continent's economic activity during the colonial era, they may have meaningfully shaped the historical trajectories of the regions that were under them. Notably, these countries are some of the world's poorest and have its highest rates of HIV prevalence. If policymakers are to remedy such global wealth and health disparities, it is important to understand whether major elements of African colonial history like these extractive institutions contributed to them, and if so, through which channels ([Nunn, 2020](#)).

However, 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 always between *different forms of extraction*, not between an extractive institution and one promoting inclusive growth (using the [Acemoglu and Robinson, 2012](#), terminology). Therefore, evaluating the comparative impacts of these regimes can sharpen our understanding of how they have shaped development and HIV prevalence.

---

<sup>1</sup> 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 the migrant-sending and mobility-restricting regimes, I focus in this paper on the sharper distinctions between these two.

**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 (“mobility restricting”). Mozambique is the southeastern country outlined in green.

The main challenge in generating such evidence is that institutions were not 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 unusual colonies like Mozambique in which it was not the only one ([Alexopoulou and Juif, 2017](#)). Located in southeastern Africa, this former Portuguese colony contained both a migrant-sending and mobility-restricting 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 mobility-restricting regime (1891-1942). Section 2 provides greater detail regarding these institutions' histories and the differences between them. According to historians, the most important one was in men's labor mobility: while it was extreme in the migrant-sending region, it was heavily restricted across the border so as to preserve a pool of low-wage and conscriptable 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 mobility-restricting 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 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 share of older men marrying younger women, and it could also have shifted parents’ preferences toward similarly-aged partners for their daughters (Irons, 1983; Leclerc-Madlala, 2008; Henrich, Boyd and Richerson, 2012). Conversely, the convergence in migration and human capital accumulation in the colonial era imply that there should not be any 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. As expected, I also 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 contributes to three main literatures. First, there are a number of influential studies examining 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 colonialist's choice from a menu of extractive institutions. My contribution in this area is to provide the first evidence on this question, establishing a chain of causality across the colonial and modern eras and establishing marriage markets as a new channel through which lasting effects arise.

Second, as transportation costs fall rapidly across the globe, it is important to understand the long-run effects of migration, both temporary and permanent (Abramitzky, Boustan and Eriksson, 2019; Derenoncourt, 2021; Dinkelman and Mariotti, 2016; Khanna, Theoharides and Yang, 2020). In the African context, there is limited evidence on the impacts of circular migration, even though it was "one of the most distinctive features of that continent's development" (Stichter, 1985, p. 1). I contribute to the literature on migration by documenting the long-run consequences of short-term labor mobility for young African men being substantially higher on one side of an arbitrary border for half a century. I also show which outcomes converge and which ones remain different after the ability to engage in circular migration had equalized.

Finally, there is an emerging literature on historical shocks as a determinant of disparities in human capital, especially health (Alsan and Wanamaker, 2018; Lowes and Montero, 2021b). Because the spread of HIV across Sub-Saharan Africa has been one of the deadliest pandemics in modern history, the spatial distribution of the virus is an important focus of studies in this area (Bertocchi and Dimico, 2019; Dwyer-Lindgren et al., 2019; Cagé and Rueda, 2020). In this respect, the most closely related paper is by Anderson (2018), who compares HIV prevalence along borders between countries with different legal regimes inherited from their European colonizers. I build on this work by providing evidence from within an African country and using the historical record to make a case for the marriage market channel.

## 2. History of the Extractive Institutions

In this section, I summarize the relevant elements of southern Mozambique's history, from the intensification of Portuguese colonization at the end of the nineteenth century to the end of its civil war and the arrival of HIV in the 1990s.

### 2.1. Assignment of Territory to Government or Company Rule

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: projecting the colonial state outward from port cities established in the sixteenth century into the surrounding regions, and granting vast, mostly unexplored areas to private companies as concessions ([Smith and Smith, 1985](#)).

Leveraging its presence in Lourenço Marques (present-day Maputo), the government assigned to itself the area from the southern international border to the Sabi River.<sup>2</sup> But it could not quickly establish state capacity north of the Sabi River to (tributaries of) the Zambezi River, so it granted a royal charter to the Mozambique Company in 1891 to govern this area ([Newitt, 1995](#)).<sup>3</sup> However, the Mozambique Company's territory was extended southward two years later. Figure 2 shows the final institutional boundaries. A royal decree defined this new southern border 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

<sup>2</sup> The government also administered a region in the north of the country around the then-capital on Mozambique Island. Because the colonial state did not establish a migrant-sending institution in the north, I omit discussion of this region for brevity. See [Newitt \(1995\)](#) for a detailed history of all of Mozambique.

<sup>3</sup> The charter was originally for 25 years but a few years later it was extended until 1942, making it the only concession in Africa to last beyond the 1920s ([Vail, 1976](#)).

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)

## *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*

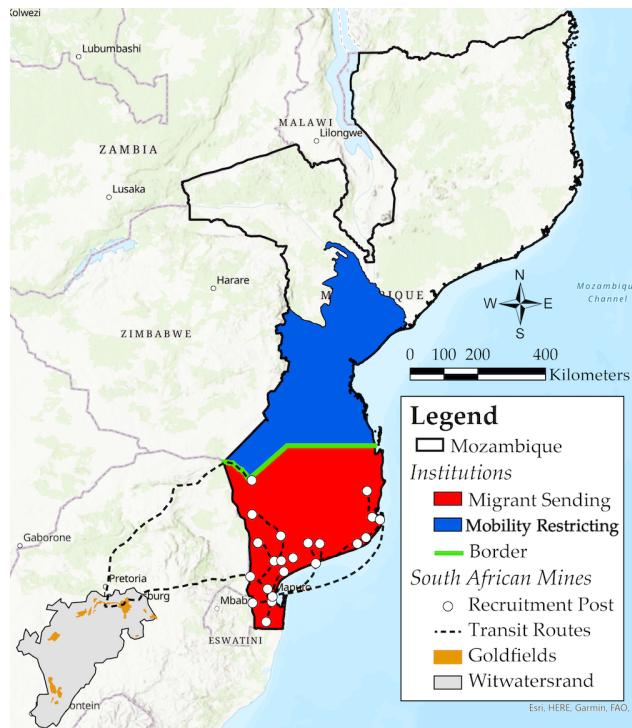
The colonial state set up a migrant-sending institution in its territory to profit from pre-existing labor flows across the border to the then-Transvaal Republic (the northeast of present-day South Africa).<sup>4</sup> The 1886 discovery of the world’s largest gold deposits on the Witwatersrand led to intense demand for African workers that men from Mozambique were vital in filling ([Clarence-Smith, 1985](#)).<sup>5</sup> 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](#)).

<sup>4</sup> 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.

<sup>5</sup> Geologists estimate that one-third of all gold ever mined is from the Witwatersrand ([Frimmel, 2019](#)).

**Figure 2:** Extractive 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 \(1946\)](#).

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](#)).<sup>6</sup> 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

<sup>6</sup> 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](#)).

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 transportation infrastructure to move them from there to the Witwatersrand. Figure 2 shows this network in 1946 and Appendix B1 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. Mobility-Restricting Institution*

In its territory, the Mozambique Company established a mobility-restricting 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 mobility-restricting 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 restricted mobility 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 with bride price customs, "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 making the payment.

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 Protestant missions established village schools in the migrant-sending region 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).<sup>7</sup> 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).<sup>8</sup>

#### *2.4. After the End of the Mobility-Restricting Institution*

The Portuguese autocrat Salazar brought about the end of the mobility-restricting institution after rising to power. He believed the Mozambique Company's concession eroded national sovereignty and decided to let it end when its royal charter was to expire (Newitt, 1995). After the colonial state took possession of the former restricted mobility region in 1942, it reorganized Mozambique's administrative boundaries. The map in Figure 3 shows the erasure of the restricted mobility region's southern boundary as the provincial border moved north to the Sabi River.

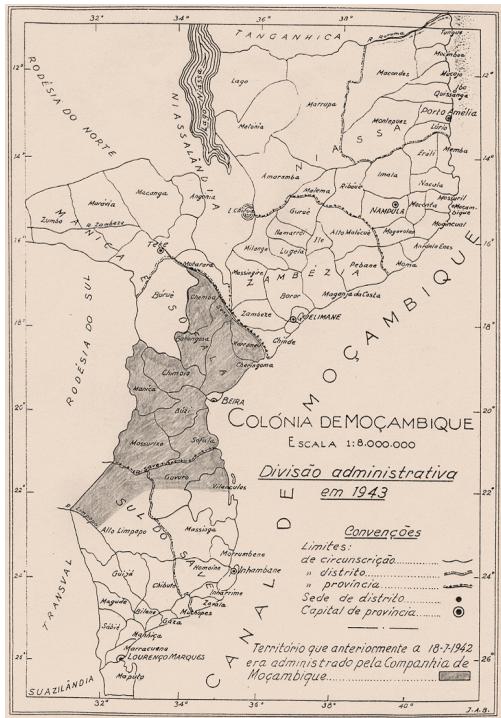
The end of the mobility-restricting institution implied much greater freedom for male labor. The visible rise in Appendix B1 in the number of Mozambicans on the Witwatersrand after 1942 is consistent with several thousand men from the restricted mobility region joining the circular migration flows each year. The Salazar regime also

---

<sup>7</sup> 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).

<sup>8</sup> 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.

**Figure 3:** Administrative Reorganization of Mozambique, 1943



Notes: Map taken from [Gengenbach \(2010\)](#) shows the administrative reorganization of Mozambique in 1943 after the Mozambique Company's concession ended. The grey area in the center of the colony is the footprint of the former restricted mobility region.

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 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 B1) and its security services aided the 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](#)).<sup>9</sup>

<sup>9</sup> 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

## 3. Colonial Data and Empirical Strategy

### 3.1. Data

To compare the impacts of the migrant-sending and mobility-restricting 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](#)). 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 mobility-restricting 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 mobility-restricting 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](#)).

#### 3.1.1. Outcomes of Interest

I focus on outcomes in three domains: labor markets, marriage and fertility, 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 marriage as the main differences between the institutions. I include the third domain because of its relationship with economic development and because colonial enumerators could measure it far more easily and accurately than income or wealth.

---

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](#)).

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.<sup>10</sup> To examine differences in marriage and fertility, I compute the ratio of ever-married men to ever-married women within a 10-year age group (“marriage ratio”) as well as the ratio of children ages 0 to 4 to women ages 15 to 44 (“child-woman ratio”).<sup>11</sup> For human capital accumulation, the outcome of interest is the share of boys and girls ages 5 to 14 enrolled in school at the time of enumeration.<sup>12</sup>

### *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 4 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.<sup>13</sup>

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

## *3.2. Empirical Strategy*

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

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

---

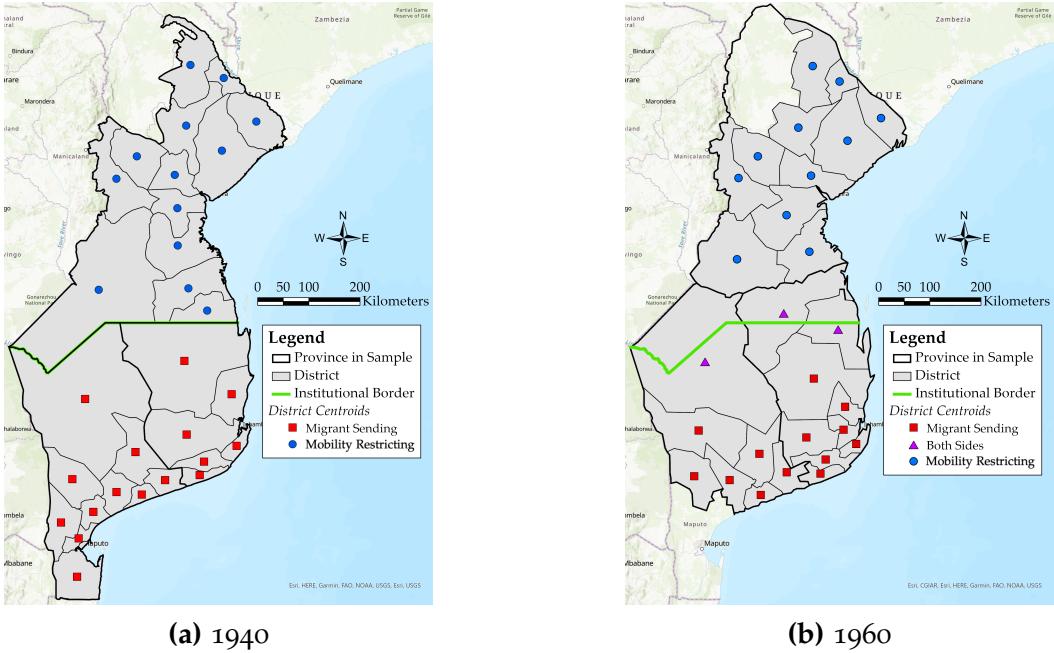
<sup>10</sup> Each census’s questionnaire asked whether a man worked abroad but the 1960 summary tables grouped circular migrants into a category with all men who worked in a mine regardless of its location. However, nearly all men in this category worked abroad, so I consider it a measure of circular migration.

<sup>11</sup> The second ratio approximates the number of children born to women of reproductive age.

<sup>12</sup> 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.

<sup>13</sup> These two cities are Lourenço Marques and Beira, the capital of the restricted mobility region.

**Figure 4: Maps of Georeferenced Colonial Census Data**



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

where  $y_d$  is the outcome of interest for district  $d$  in the set  $B$  defined by the bandwidth restrictions above.<sup>14</sup> The explanatory variables are  $\text{MigrantSending}_d$ , 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  $\text{Lon}_d$ , a centroid's longitude coordinate, which Kelly (2020) recommends including in RD designs to capture east-west trends.<sup>15</sup> I use a local linear specification estimated 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  $\tau$  in equation (1) identifies the effect of historical assignment to the migrant-sending institution *relative to historical assignment to the mobility-restricting institution*. The motivating idea is that because the border between them was arbitrary,

<sup>14</sup> Because the colonial data are reported at the district level, there are too few centroids near the border to estimate the Calonico, Cattaneo and Titiunik (2014) mean squared error (MSE) optimal bandwidth.

<sup>15</sup>  $\text{Distance}_d$  has a near-perfect correlation with latitude ( $\rho > 0.99$ ), so it accounts for north-south trends.

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  $\tau$  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.

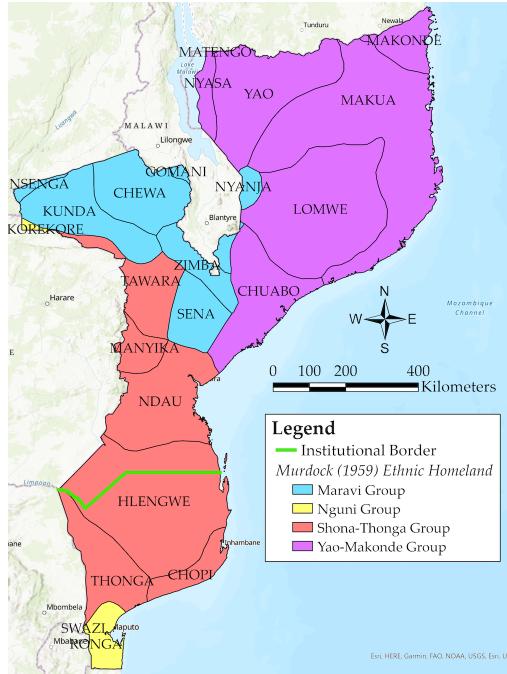
An important concern when conducting inference in geographic RD designs is positive spatial autocorrelation (Kelly, 2020). Due to the relatively small area under examination and the slow rate at which many outcomes change across space, estimated standard errors may be too small due to similarity among 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 Conley 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”) (Conley, 1999; Colella et al., 2020). I report these standard errors in addition to those robust to heteroskedasticity. 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.<sup>16</sup> 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

---

<sup>16</sup> 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).

**Figure 5:** Ethnic Group Homelands in Mozambique



Notes: Map shows Murdock (1959) ethnic homelands by cultural group. The green line is the institutional border.

between  $-1$  and  $1$ .<sup>17</sup> I report the difference of the observed and expected  $I$ -statistics as the measure of spatial autocorrelation along with  $I$ 's standard deviation.<sup>18</sup>

### 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 5 shows that the border is entirely within one Murdock (1959) ethnic homeland. Additionally, the neighboring ethnicities are all part of the Tsonga 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  $0.25 \times 0.25$  degree cells—approximately  $25 \text{ km} \times$

<sup>17</sup> The resulting bandwidth is approximately 170 km for 1940 and approximately 140 km for 1960.

<sup>18</sup> The asymptotic distribution of the  $I$ -statistic is standard normal.

**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 $p$	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 Autocorrelation SD	0.02	0.03	0.02	0.02	0.02	0.02
Mobility Restricting Mean	182.7	40.94	0.179	49.19	10.84	1.259

*Notes:* Observations are  $0.25 \times 0.25$  degree cells. Standard errors clustered by administrative post are in parentheses and Conley standard errors using a 100-km bandwidth and a Bartlett kernel are in brackets. 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 chosen to minimize mean squared error, as suggested by ([Calonico, Cattaneo and Titiunik, 2014](#)). Data sources and variable definitions are in Appendix A.

25 km in the study area—and estimate equation (1) clustering standard errors by third-level administrative unit (“administrative posts,” shown in Appendix B2).<sup>19</sup> Consistent with the border being arbitrary, Table 1 shows that changes in these variables just inside the migrant-sending institution are small relative to restricted mobility means.

#### 4. Effects of the Institutions during the Colonial Era

I now turn to studying the institutions’ impacts on labor markets, marriage and fertility, and human capital accumulation while Mozambique was a Portuguese colony. Table 2 reports the RD estimates for each outcome of interest two years before the abolition of the mobility-restricting institution (Panel A) and 18 years after (Panel B), and Figure 6 presents RD plots for selected outcomes in each year. The results show that the institutions differed substantially in men’s circular migration and marriage and fertility while they both existed, but only marriage market outcomes remained markedly different nearly two decades after the end of the mobility-restricting regime.

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

#### ***4.1. Labor Markets***

Due to historians' emphasis on circular migration as the main difference between the institutions, I first examine whether it changed at the border in 1940. 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 across the border worked abroad. Figure 6a shows that only in three restricted mobility districts did any men work abroad.

There is important heterogeneity in working abroad by age. Appendix B3 shows that rates of circular migration were highest among ages 15 to 24 and 25 to 34 (around 40 percent for both groups) as well as a significant decline among older groups. This pattern is consistent with historical narratives regarding young men's motivations for working abroad and the requisite physical fitness required to do so.

However, Panel B Column (1) shows convergence in men's circular migration after the mobility-restricting institution ended, when 16 percent of prime-aged men there worked abroad.<sup>20</sup> This pattern implies that the regime significantly constrained men's mobility and its end in 1942 led to major changes in their occupational choices.

Next, I examine whether differences between the institutions affected women's occupations. In spite of men's absences, the estimates in Column (2) in Panels A and B show no difference at the border in the near-universal share of prime-aged women working in agriculture. These results suggest that men's absences did not affect women's responsibility for food production. Indeed, this division of labor likely predicated the institutions and enabled men's circular migration in the first place.

#### ***4.2. Marriage and Fertility***

I then turn to comparing marriage market outcomes across the institutional border, as narrative histories argue that earning bridewealth was an important motivation for men

---

<sup>20</sup> I exclude a migrant-sending district with only one circular migrant, which is most likely an error.

**Table 2:** Comparing Institutions during the Colonial Era

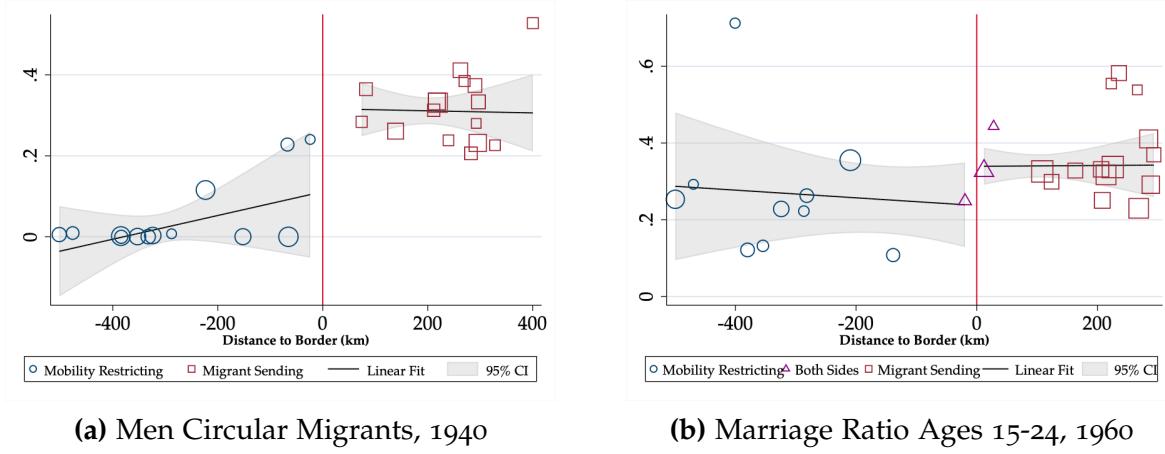
Panel A: 2 Years before End of Mobility Restricting Institution (1940)							
	Labor Markets		Marriage and Fertility Ratios			Human Capital	
	Men Migrants (1)	Women Farming (2)	Marriage Ages 15-24 (3)	Marriage Ages 25-34 (4)	Child- Woman (5)	Boys in School (6)	Girls in School (7)
Migrant Sending	0.207 (0.089) [0.085]	0.015 (0.018) [0.016]	0.269 (0.080) [0.061]	0.312 (0.113) [0.088]	0.202 (0.098) [0.092]	-0.028 (0.016) [0.015]	-0.001 (0.006) [0.006]
Observations	29	29	28	28	28	29	29
Bandwidth	-503, 401	-503, 401	-503, 401	-503, 401	-503, 401	-503, 401	-503, 401
Spatial Autocorrelation	-0.14	-0.10	0.26	0.14	-0.15	-0.03	-0.06
Spatial Autocorrelation SD	0.11	0.10	0.11	0.11	0.12	0.11	0.10
Mobility Restricting Mean	0.047	0.958	0.364	0.696	0.848	0.050	0.006
Panel B: 18 Years after End of Mobility Restricting Institution (1960)							
	Labor Markets		Marriage and Fertility Ratios			Human Capital	
	Men Migrants (1)	Women Farming (2)	Marriage Ages 15-24 (3)	Marriage Ages 25-34 (4)	Child- Woman (5)	Boys in School (6)	Girls in School (7)
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]	0.056 (0.086) [0.082]	-0.014 (0.033) [0.029]	-0.026 (0.024) [0.021]
Observations	27	28	27	27	28	28	28
Bandwidth	-500, 294	-500, 294	-500, 294	-500, 294	-500, 294	-500, 294	-500, 294
Spatial Autocorrelation	-0.09	-0.02	-0.03	-0.15	0.05	0.44	0.24
Spatial Autocorrelation SD	0.12	0.11	0.12	0.10	0.12	0.12	0.13
Mobility Restricting Mean	0.163	0.997	0.267	0.635	0.807	0.089	0.041

*Notes:* Observations are districts. Robust standard errors are in parentheses and Conley standard errors using a 100-km bandwidth and a Bartlett kernel are in brackets. Regressions estimate a local linear RD specification on each side of the border using a triangular kernel and include longitude as a control. The left (negative) and right (positive) ends of the RD bandwidth used in each panel are in kilometers. Data sources, variable definitions, and RD bandwidth selection criteria are in Section 3. Columns (3) and (4) in both panels as well as Panel A Column (5) and Panel B Column (1) each exclude an extreme outlier for that outcome as described in the text.

to work abroad. I focus on the 10-year age groups most heavily affected by circular migration (ages 15 to 24 and 25 to 34), as its marriage market impacts should have been most apparent for them. I also examine fertility because its response to circular migration could have been in either direction—decreasing if it led parents to focus on child quality or increasing if more children raised subsistence agricultural output.

Table 2 Panel A Columns (3) and (4) show that while the two institutions still

**Figure 6:** 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.

existed, the marriage ratio for both age groups increased by around 0.3 just inside the migrant-sending region.<sup>21</sup> The point estimates are large relative to ratios in the mobility-restricting institution for ages 15 to 24 (averaging around 1:3) and 25 to 34 (2:3). Panel B Columns (3) and (4) show that these differences decreased but remained sizable after the mobility-restricting regime ended, though the estimate for ages 25 to 34 is imprecise. The RD plot in Figure 6b suggests 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. I discuss why marriage differences could continue in Section 5.

Fertility also was higher on the migrant-sending side of the border in 1940. The estimate in Panel A Column (5) is of a 0.2-point increase in the child-woman ratio just inside the migrant-sending region, which is also large relative to the mean of 0.85 in restricted mobility districts.<sup>22</sup> The positive response of fertility to circular migration (and its higher wages) suggests that Malthusian dynamics were at work, which is

<sup>21</sup> These columns exclude a district in the mobility-restricting 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 almost certainly an error given that it is nearly 3 times the average ratio among other districts in the region.

<sup>22</sup> The estimate excludes a district just inside the migrant-sending region with a child-woman ratio of 1.3, far higher than any other district in the sample. Its exclusion makes the estimated coefficient smaller.

unsurprising given the widespread reliance on subsistence agriculture.

The fertility estimate for 1960 in Panel B Column (5) is smaller than the one for 1940 and it is imprecise. The reduction in size is consistent with fertility responding positively to increased rates of circular migration in the former restricted mobility region. Because marriage was generally a prerequisite for having children in this period, it is also consistent with declining but still meaningful differences in marriage at the border.

#### *4.3. Human Capital Accumulation*

Finally, I examine whether differences between the institutions affected investments in children's human capital. Table 2 Panel A Column (6) shows that prior to the end of the mobility-restricting institution—after Protestant missionaries were expelled from the migrant-sending region but before Catholic missions assumed responsibility for education there—boys just inside the migrant-sending region were 3 p.p. less likely to be in school. This effect is large relative to the 5-percent enrollment rate for boys in the mobility-restricting institution. In contrast, there is no effect for girls in Column (7).

However, after the mobility-restricting regime ended and Catholic missions became responsible for education throughout Mozambique, the difference in boys' schooling in Panel B Column (6) became smaller and imprecise. It is difficult to argue whether the convergence in schooling provision or that in circular migration mattered more, but clearly their combination eroded differences in enrollment for boys. There is a larger but imprecise coefficient for girls in Column (7), though the absence of a difference in 1940 suggests that the emergence of one in 1960 was not due to institutional differences.

### **5. Modern Data and Empirical Strategy**

In this section, I lay out a conceptual framework that explains why colonial-era differences in marriage markets should continue to today despite the convergence in circular migration. I also discuss the modern data I use as well as refinements of the

RD estimation in the previous section.

### *5.1. Conceptual Framework Linking Past and Present*

The colonial-era results—more men marrying similarly-aged women in the migrant-sending region, no lasting differences in men’s circular migration or schooling—have clear implications for HIV prevalence and economic development today. Specifically, seroprevalence should be lower just inside the migrant-sending region while there should not be any differences in wealth levels along the border.

#### *5.1.1. HIV Prevalence*

With respect to HIV, smaller age gaps between partners in sub-Saharan Africa lower the risk of contracting the virus ([Schaefer et al., 2017](#)). Intuitively, older men transmit it to younger women, who as they age transmit it to men of similar ages, perpetuating the cycle ([de Oliveira et al., 2017](#)). An implication is that fewer age-disparate relationships should lower HIV prevalence and give its age profile a later peak, especially for women.

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 that they have continued is the additional decades of greater marriageability for young men in that region. In polygynous marriage markets with bride price customs, many young women marry men of an older generation, reducing the pool of potential matches for young men. As they age, these single men marry young women because those her age are the only ones still unmarried, perpetuating the cycle of age-disparate relationships.

But in the former migrant-sending region, nearly a century of circular migration—and thus greater marriageability for young men—may have dampened this pattern. With each successive generation of young men marrying more similarly-aged women, the share of older men with young wives should have decreased, enabling

more young men in subsequent generations to marry young women. In contrast, because there was only 35 years of circular migration from the former restricted mobility region, it should not have experienced the same reduction in age disparities.

In addition to these mechanical reasons for continued marriage and dating market differences, Leclerc-Madlala's discussion of marriage, dating, and HIV in Southern Africa suggests social norms may also play a role:

Girls are encouraged to seek older men as husbands and are often cautioned against marrying a man of similar age, as this is believed to lead to marriage instability. ... [Age-disparate relationships] have 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. ([Leclerc-Madlala, 2008](#), pp. S22-S23)

Applying this anthropological framework to the context of southern Mozambique, it is possible that the longer shock to young men's marriageability in the former migrant-sending region did more to weaken social norms about age-disparate relationships there. The intuition is that increasing the number of parents married to similarly-aged spouses should have reduced the belief that such marriages are unstable or even taboo. As a result, there should be more of these partnerships in the next generation.

The equalization of circular migration in the colonial era and South Africa's severe restrictions on it after Mozambique's independence also have important implications for HIV prevalence. Converging rates of circular migration should have equalized the risks of transmission associated with this phenomenon along the border ([Weine and Kashuba, 2012](#)). In addition, the sharp reduction in it a decade before HIV exploded across Southern Africa would also delay the virus's arrival in this area, as would mobility limitations during the 1977-92 civil war ([Iliffe, 2006](#); [Audet et al., 2010](#)).

### *5.1.2. Economic Development*

In contrast, the colonial-era patterns imply that there should not be any differences in levels of economic development today. Convergence in circular migration rates also should have reduced any differences in wages, as it allowed families on both sides of the border to benefit from circular migration (Khanna, Theoharides and Yang, 2020). Moreover, the convergence in school enrollment rates for boys implies that there is no reason related to the institutions for human capital accumulation to be different.

## *5.2. Data*

To test these predictions regarding present-day HIV prevalence and economic development, I use georeferenced individual-level data from the 2009, 2011, 2015, and 2018 waves of the Demographic and Health Surveys (DHS) in Mozambique. Figure 7 shows the reported locations of the survey clusters within 200 km of the institutional border. These locations are slightly displaced to protect the anonymity and privacy of respondents.<sup>23</sup>

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 D, 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.

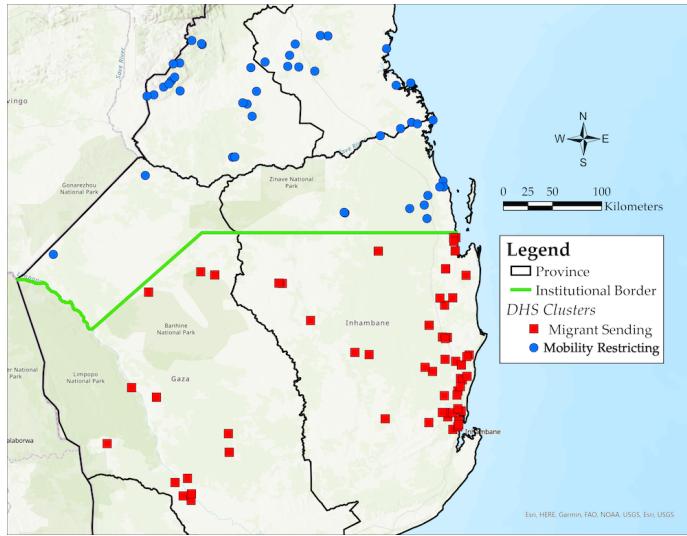
The outcome of interest when examining HIV is the result of blood tests for the virus from a random subset of respondents in 2009 and 2015. I restrict this analysis to adults ages 15 to 64.<sup>24</sup> For economic development, the variables of interest are an index of household asset ownership (measured in 2009, 2011, 2015, and 2018), an indicator for

---

<sup>23</sup> 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.

<sup>24</sup> 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 7: Map of Georeferenced DHS Clusters**



*Notes:* Map shows the reported locations of survey clusters within 200 km of the border in the 2009, 2011, 2015, and 2018 DHS waves in Mozambique. The shapefiles are from [Ministério de Saúde, Instituto Nacional de Estatística and ICF Macro \(2010\)](#), [Ministério de Saúde, Instituto Nacional de Estatística and ICF International \(2013\)](#), [Ministério de Saúde, Instituto Nacional de Estatística and ICF \(2018\)](#), and [Instituto Nacional de Estatística and ICF \(2019\)](#).

whether a child is stunted (2011), and years of schooling (2009, 2011, and 2015).<sup>25</sup>

### 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 mobility-restricting 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  $\delta_t$ .

<sup>25</sup> The index equals a household's quintile in the first principal component of a principal component analysis of its assets (1 = lowest, 5 = highest). Children are considered stunted if their height-for-age z-scores using the World Health Organization's Child Growth Standards are less than -2.

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  $B_{\text{MSE}}^*$ . I continue to use a local linear RD specification with a triangular kernel.

### 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 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. 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 again I address by calculating Conley standard errors and Moran  $I$ -statistics as discussed in Section [3.2.1](#).<sup>26</sup>

## 6. Effects of the Former Institutions in the Modern Era

In this section, I study the institutions’ impacts on HIV prevalence and economic development today. Tables [3](#) and [4](#) report RD estimates for the respective outcomes, and Figure [8](#) presents graphical evidence on seroprevalence. These results show that, 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

---

<sup>26</sup> 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
Mobility Restricting 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](#)).

10 p.p. just inside the migrant-sending institution. This point estimate is large relative to the 22 percent of the mobility-restricting 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.

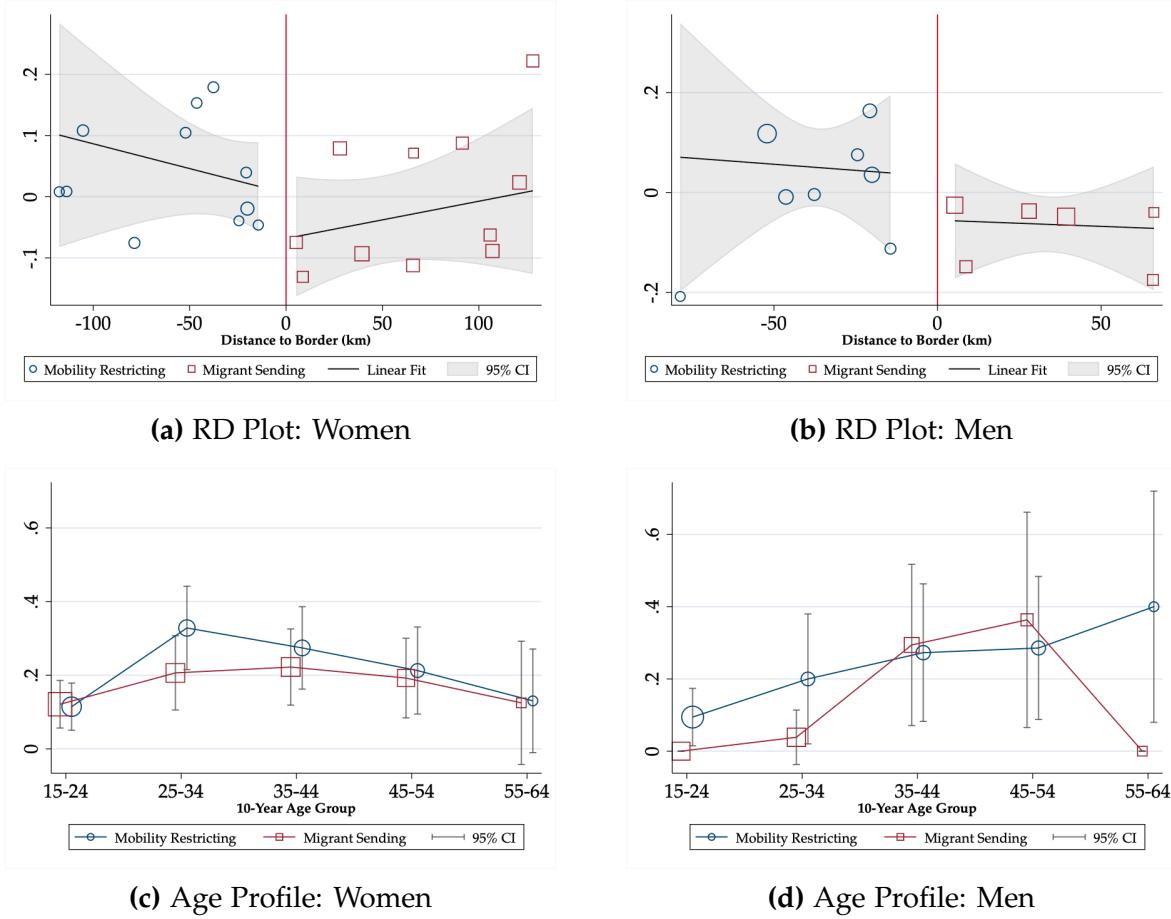
To probe the robustness of this result, I split the sample by sex in Columns (2) and (3), which shows that this effect is of equal magnitude for women and men. Figures 8a and 8b provide visual evidence of these sizable differences at the border. However, after accounting for the number of clusters the male estimate is imprecise, likely as a result of the much smaller sample size. I also show in Appendix C1 that there are no substantive differences in blood test refusal rates along the border, helping to rule out selection into testing due to history or other factors ([Lowes and Montero, 2021b](#)).

To rationalize the large effect sizes estimated above, I compare the age profiles of HIV prevalence in the two institutions.<sup>27</sup> Specifically, I calculate the mean seroprevalence for

---

<sup>27</sup> 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](#)).

**Figure 8: HIV Prevalence RD Plots and Age Profiles**



*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. Age profiles show the mean HIV prevalence for each sex within a 10-year age group within the MSE-optimal RD bandwidth in the two institutions. Shape sizes in both plot types reflect the relative number of adults in a cluster or age group.

each 10-year age group within each sex's MSE-optimal RD bandwidth.<sup>28</sup> Figures 8c and 8d plot 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 restricted mobility region. It is most apparent for women ages 25 to 34, when HIV prevalence peaks in the former mobility-restricting institution. There is a similar pattern for men but the small sample size results in substantial noise.

These age profiles of HIV prevalence are consistent with the [de Oliveira et al. \(2017\)](#)

<sup>28</sup> The HIV sample is too small to permit reliable RD estimation by sex and age group.

**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
Mobility Restricting 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](#)).

transmission cycle and its implications discussed in Section 5.1. The peak of women's seroprevalence in the former mobility-restricting institution is both larger than the one across the border and farther away from the male peak. This pattern could arise from relationships with wider age disparities, a channel I examine 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 mobility-restricting institution. The coefficients for childhood stunting (-6 p.p.) and female schooling (0.38 years) are non-trivial relative to the respective restricted mobility means (38 percent and 2.5 years) but notably imprecise.

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 mobility-restricting institution's abolition have disappeared in the intervening decades.

## 7. Explaining Differences in HIV Prevalence

I now study the channels that have led 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 examine other important HIV risk factors, splitting them into ones the public health literature links to age gaps and those that are unrelated. The main result is that age gaps between spouses and sexual partners are markedly smaller and behavioral risk factors associated with them are less common just inside the former migrant-sending institution. They are the primary differences in HIV risk at the border, suggesting colonial-era patterns substantially affect current seroprevalence.

### 7.1. Age-Disparate Relationships

To examine age disparities between spouses and sexual partners, I use data from the IPUMS 10-percent sample of the 2007 Mozambican census in addition to the DHS. The former allow 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.<sup>29</sup> However, the census lacks information on sexual activity and the data are at a much coarser geographic resolution (administrative posts, see Appendix B2 for a map).<sup>30</sup>

---

<sup>29</sup> 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.

<sup>30</sup> 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.

**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
Mobility Restricting Mean	8.380	7.265	5.110

*Notes:* Standard errors clustered by administrative post (census data) or survey cluster (DHS data) 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 Titunik, 2014](#)).

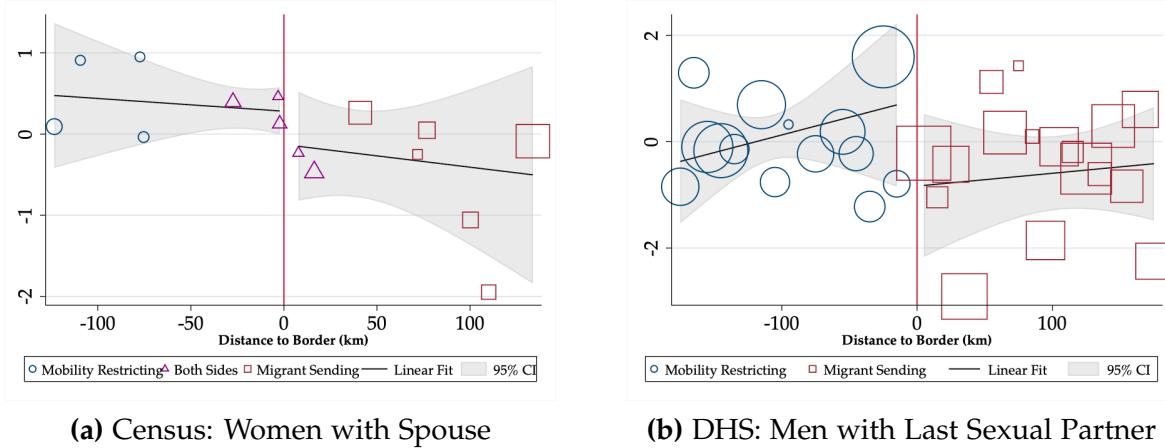
The outcome of interest in both datasets is the man's age minus the woman's, which I winsorize at 90 percent due to extreme outliers at both ends of the distribution.<sup>31</sup> I examine age gaps between women of any age and their spouses in the census sample to make the closest link possible between colonial-era and present-day marriage market outcomes. To connect them to sexual behavior, I also study age gaps between adults in the DHS ages 15 to 49 ("reproductive age") and their most recent sexual partner.<sup>32</sup>

Table 5 reports the results of estimating equation (2) for age disparities in these datasets and Figure 9 presents RD plots for these outcomes. Column (1) shows that the age disparity between a woman and her linked spouse in the census is 0.8 years smaller just inside the former migrant-sending institution. This estimate is meaningful relative to the average of 8.4 years in the former restricted mobility region, and its precision does

<sup>31</sup>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. The 90-percent winsorization increases the precision of the RD estimate without changing its magnitude.

<sup>32</sup>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.

**Figure 9: RD Plots for Age-Disparate Relationships**



*Notes:* RD plots show the mean age disparity among the specified group in an administrative post (census data) or DHS survey cluster within a 10-km bin (due to the high number of clusters), net of age, age squared, longitude, and year fixed effects. The running variable is 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 an administrative post or bin.

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 between sexual partners in the DHS. These estimates for women (-3.1 years) and men (-1.9 years) are even larger relative to the respective restricted mobility means (7.3 years and 5.1 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. Taken together, these results suggest age-disparate relationships are important in explaining the HIV prevalence result.

## 7.2. Risk Factors Associated with Age-Disparate Partnerships

While age gaps in relationships can be HIV risk factors on their own, they are also associated with behaviors facilitating transmission of the virus. They include male partners who are in concurrent relationships, an earlier sexual debut for women and girls, and not using condoms (Evans et al., 2019; Mabaso et al., 2021; Schaefer et al., 2017). I measure these outcomes among reproductive-age adults in the DHS.

Table 6 Column (1) shows that men in the former migrant-sending region are 16

**Table 6:** HIV Risk Factors Associated with Age Disparities

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

*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](#)).

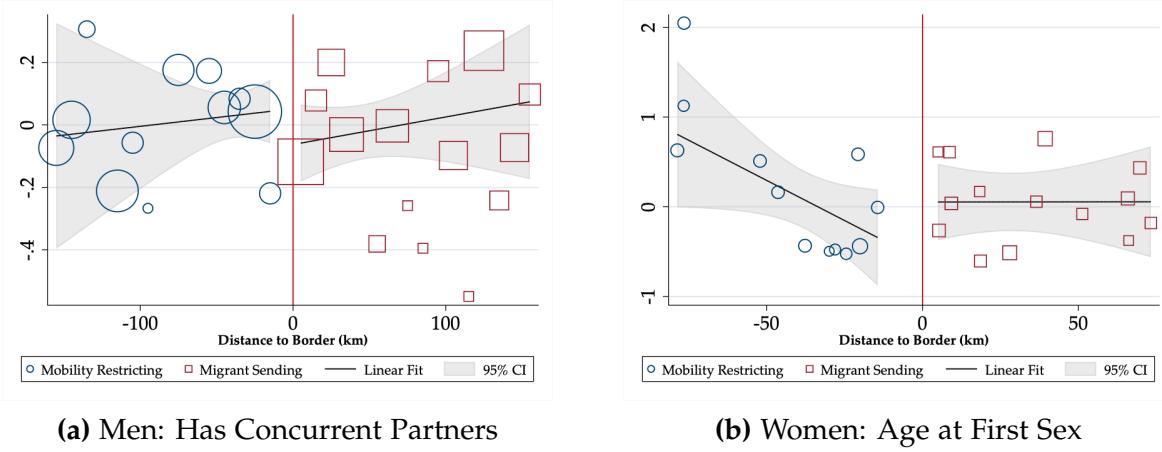
p.p. less likely to have concurrent sexual partners, which substantial compared to the restricted mobility mean of 26 percent. In Column (2), I estimate that girls' sexual debuts occur 0.8 years later in the former migrant-sending region, which is non-trivial relative the restricted mobility mean of 16.1 years. Although the estimates are somewhat imprecise, the changes at the border are apparent in Figure 10, especially for concurrency.

In contrast, the evidence on condom use in the most recent sexual intercourse is less compelling. The coefficient for men in Column (3) is large—6.5 p.p. relative to a restricted mobility mean of 7 percent—but notably imprecise, and in Column (3) there is a null RD estimate for women. Nonetheless, the direction of the results in Table 6 is consistent with reductions in risk factors associated with age-disparate relationships.

### 7.3. Other HIV Risk Factors

Finally, I examine 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

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



Notes: RD plots show the mean of the specified outcome within a 10-km bin (left panel, due to the high number of clusters) or in DHS survey clusters (right panel), 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.

continent, I create indicator variables for: having a genital ulcer in the past 12 months ([Chen et al., 2000](#)), a polygynous marriage ([Reniers and Tfaily, 2008](#)), a woman having ever been forced to engage in sexual activity, a man having ever paid for sex ([Dunkle et al., 2004](#)), a woman being able to decide alone on her own healthcare ([Anderson, 2018](#)), and a man having been medically circumcised ([Maffioli, 2017](#)).

In Appendix C2, I estimate 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 restricted mobility mean of 8.7 percent) and more men who are medically circumcised (3.1 p.p. relative to a restricted mobility 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 compared to a restricted mobility mean of 0.5 percent—*increases* HIV risk in the former migrant-sending region. As such, these results suggest that the commonly-cited risk factors above do not explain the drop in HIV prevalence just inside the former migrant-sending area.

## 8. Conclusion

This paper show that two institutions common throughout colonial Africa can have markedly different impacts on HIV prevalence today but result in no long-run differences in economic development. The effects on HIV appear to arise through the institutions' lasting effects on marriage markets: where men historically were able to marry women of more similar ages, seroprevalence is much lower. The continuation of these differences in marriage and dating from the colonial era to today is plausible given the length of the shock to marriageability for young men—nearly a century—and the overlapping generations nature of these markets.

The results in this paper speak to the importance of history and institutions in shaping present-day disparities in the HIV pandemic and document the marriage market as a novel channel through which their lasting effects can arise. Such lessons are important for policymakers to take into account, as effectively combatting global health disparities in one of the modern era's deadliest pandemics requires a full understanding of its historical and social roots. These findings also provide insight into the long-run consequences of migration on health and wealth, which are important to understand given how rapidly transportation costs have fallen—and continue to fall—around the world in this era of globalization.

## References

- Abramitzky, Ran, Leah Boustan, and Katherine Eriksson.** 2019. "To the New World and Back Again: Return Migrants in the Age of Mass Migration." *Industrial and Labor Relations Review*, 72(2): 300–322. [5]
- 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.** 2015. "The Effect of the Tsetse Fly on African Development." *American Economic Review*, 105(1): 382–410. [45]
- 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, 35]
- Ashraf, Quamrul, and Oded Galor.** 2011. "Dynamics and Stagnation in the Malthusian Epoch." *American Economic Review*, 101(5): 2003–2041. []
- 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, 24]
- 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. [13]

- 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. [15, 18, 27, 28, 30, 32, 34, 48, 49]
- 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. [27]
- Cattaneo, Matias D., Nicolás Idrobo, and Rocío Titiunik.** 2019. *A Practical Introduction to Regression Discontinuity Designs: Foundations*. Cambridge, UK:Cambridge University Press. [15]
- Chen, Cheng Y., Ronald C. Ballard, Consuelo Beck-Sague, Yusuf Dangor, Frans Radebe, Scott Schmid, Judith Weiss, Vaness Tshabalala, Glenda Fehler, Ye Htun, and Stephen Morse.** 2000. "Human Immunodeficiency Virus Infection and Genital Ulcer Disease in South Africa." *Sexually Transmitted Diseases*, 27(1): 21–29. [35]
- Chiappori, Pierre-André, Murat Iyigun, and Yoram Weiss.** 2009. "Investment in Schooling and the Marriage Market." *American Economic Review*, 99(5): 1689–1713. []
- Clarence-Smith, Gervase.** 1985. *The Third Portuguese Empire, 1825–1975: A Study in Economic Imperialism*. Manchester:Manchester University Press. [7, 8]
- Colella, Fabrizio, Rafael Lalive, Seyhun Orcan Sakalli, and Mathias Thoenig.** 2020. "Inference with Arbitrary Clustering." [16]
- Conley, T.G.** 1999. "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics*, 92(1): 1–45. [16]
- 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. [46]
- Danielson, Jeffrey J., and Dean B. Gesch.** 2011. "Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010)." U.S. Geological Survey Open-File Report 2011-1073, Reston, VA. [45]
- Darch, Colin.** 1983. "Notas sobre Fontes Estatísticas Oficiais Referente à Economia Colonial Moçambicana: Uma Crítica Geral." *Estudos Moçambicanos*, 4: 103–125. [13]
- 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. [5]

**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. [4, 23, 29]

**Derenoncourt, Ellora.** 2021. "Can You Move to Opportunity? Evidence from the Great Migration." [5]

**Dinkelman, 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, 5]

**Direcção Provincial dos Serviços de Estatística.** 1966. *III Recenseamento Geral da População na Província de Moçambique*. Lourenço Marques:Instituto Nacional de Estatística. [13]

**Dunkle, Kristin L., Rachel K. Jewkes, Heather C. Brown, Glenda E. Gray, James A. McIntryre, and Siobán D. Harlowa.** 2004. "Transactional Sex among Women in Soweto, South Africa: Prevalence, Risk Factors and Association with HIV Infection." *Social Science & Medicine*, 59(8): 1581–1592. [35]

**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, and Jason B. Hall, 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, 33]

**Fernández Cebrián, Pablo.** 2021. "Withdrawal of the State: The Provision of Primary Schooling in Mozambique under the Indigenato." [12]

**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. [7]

**Fuller, Charles Edward.** 1955. "An Ethnohistoric Study of Continuity and Change in Gwambe Culture." PhD diss. Northwestern University, Evanston, IL. [11]

**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. [12]

**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. [7]

**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. [12]

**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]

**Hansen, Gary D., and Edward C. Prescott.** 2002. "Malthus to Solow." *American Economic Review*, 94(2): 1205–1217. []

**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. [13]

**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. [13]

**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. [13, 24]

**Instituto Nacional de Estatística, and ICF.** 2019. *Moçambique Inquérito Nacional sobre Indicadores de Malária*. Rockville, MD. [26]

- 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. [12]
- Junod, Henri A.** 1912. *The Life of a South African Tribe*. Vol. 1, London:David Nutt. [7, 10]
- Kelly, Morgan.** 2020. "Understanding Persistence." [15, 16]
- Khanna, Gaurav, Caroline Theoharides, and Dean Yang.** 2020. "Abundance from Abroad: Migrant Income and Long-Run Economic Development." [5, 25]
- Kiszewski, Anthony, Andrew Mellinger, Andrew Spielman, Pia Malaney, Sonia Ehrlich Sachs, and Jeffrey Sachs.** 2004. "A Global Index Representing the Stability of Malaria Transmission." *American Journal of Tropical Medicine and Hygiene*, 70(5): 486–498. [45]
- 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." [5]
- Lowes, Sara, and Eduardo Montero.** 2021b. "The Legacy of Colonial Medicine in Central Africa." *American Economic Review*, 111(4): 1284–1314. [5, 28]
- 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, 33]
- Maffioli, Elisa M.** 2017. "Is Traditional Male Circumcision Effective as an HIV Prevention Strategy? Evidence from Lesotho." *PLOS ONE*, 12(5): e0177076. [35]
- Michalopoulos, Stelios, and Elias Papaioannou.** 2014. "National Institutions and Subnational Development." *Quarterly Journal of Economics*, 129(1): 151–213. [5]
- Michalopoulos, Stelios, and Elias Papaioannou.** 2016. "The Long-Run Effects of the Scramble for Africa." *American Economic Review*, 106(7): 1802–1848. [5]
- Ministério de Saúde, Instituto Nacional de Estatística, and ICF.** 2018. *Moçambique Inquérito de Indicadores de Imunização, Malária e HIV/SIDA 2015*. Maputo. [26]
- Ministério de Saúde, Instituto Nacional de Estatística, and ICF International.** 2013. *Moçambique Inquérito Demográfico e de Saúde 2011*. Calverton, MD. [26]

**Ministério de Saúde, Instituto Nacional de Estatística, and ICF Macro.** 2010. *Inquérito Nacional de Prevalência, Riscos Comportamentais e Informação sobre o HIV e SIDA em Moçambique 2009*. Calverton, MD. [26]

**Ministério do Ultramar.** 1959. "Carta de Moçambique." [14]

**Minnesota Population Center.** 2020. *Integrated Public Use Microdata Series, International: Version 7.3*. Minneapolis, MN:IPUMS. [46]

**Moran, P. A. P.** 1950. "Notes on Continuous Stochastic Phenomena." *Biometrika*, 37(1-2): 17–23. [16]

**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*. [50]

**Mozambique News Agency.** 2000. "Tourism Brings in \$32 Million." *AIM Reports*. [50]

**Mozambique News Agency.** 2011. "President Inaugurates New Terminal at Vilankulo Airport." *AIM Reports*. [50]

**Murdock, George Peter.** 1959. *Africa: Its Peoples and Their Culture History*. New York:McGraw-Hill Book Company. [17]

**Méndez-Chacón, Esteban, and Diana Van Patten.** 2021. "Multinationals, Monopsony, and Local Development: Evidence from the United Fruit Company." [5]

**Newitt, Malyn.** 1995. *A History of Mozambique*. Bloomington, IN:University of Indiana Press. [6, 7, 8, 11]

**Nunn, Nathan.** 2020. "The Historical Roots of Economic Development." *Science*, 367(6485): eaaz9986. [1]

**Portugal.** 1900. *Collecção Official de Legislação Portugueza: Anno de 1899*. Lisboa:Imprensa Nacional. [9]

**Reniers, Georges, and Rania Tfaily.** 2008. "Polygyny and HIV in Malawi." *Demographic Research*, 19(53): 1811–1830. [35]

**Repartição Nacional de Estatística.** 1942. *Censo da População em 1940*. Vol. 2-5, Lourenço Marques:Imprensa Nacional de Moçambique. [13]

**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." [14]
- 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. [4, 23, 33]
- Schneider, Udo, Andreas Becker, Peter Finger, Elke Rustemeier, and Markus Ziese.** 2020. *GPCC Full Data Monthly Product Version 2020 at 0.25°: Monthly Land-Surface Precipitation from Rain-Gauges Built on GTS-Based and Historical Data*. Global Precipitation Climatology Center. [45]
- 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. [6]
- 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." [8]
- 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. [28]
- 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]
- World Bank.** 2020. *World Slope Model in Degrees*. The World Bank Group. [45]
- Yang, Dean.** 2008. "International Migration, Remittances and Household Investment: Evidence from Philippine Migrants' Exchange Rate Shocks." *Economic Journal*, 118(528): 591–630. [3]

**Zabel, Florian, Birgitta Putzenlechner, and Wolfram Mauser.** 2014. "Global Agricultural Land Resources: A High Resolution Suitability Evaluation and Its Perspectives until 2100 under Climate Change Conditions." *PLOS ONE*, 9(12): e114980. [45]

## Appendix A. Data Sources and Variable Definitions

### Geographic Traits [18]

- *Elevation*: Average altitude in meters in the  $0.25 \times 0.25$  degree cell. Data from [Danielson and Gesch \(2011\)](#).
- *Rainfall*: Average precipitation in millimeters in the  $0.25 \times 0.25$  degree cell from 1891 to 2016. Data from [Schneider et al. \(2020\)](#).
- *Slope*: Average slope in degrees in the  $0.25 \times 0.25$  degree cell. Data from [World Bank \(2020\)](#).
- *Soil Index*: Average agricultural suitability index value for growing 16 food and energy crops from 1981 to 2010 in the  $0.25 \times 0.25$  degree cell. Data from [Zabel, Putzenlechner and Mauser \(2014\)](#).

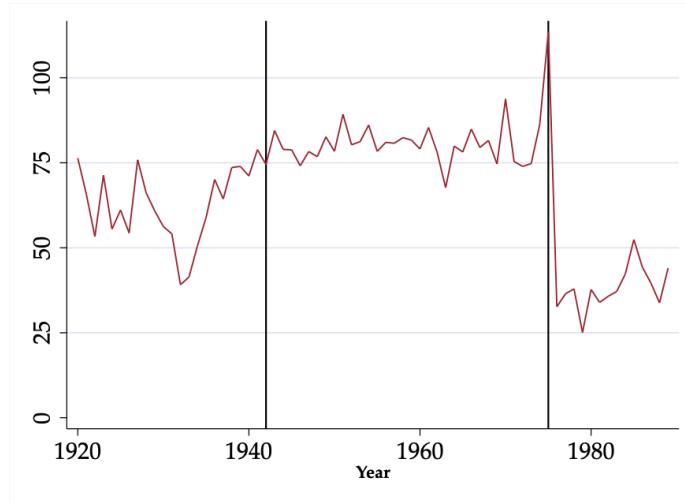
### Disease Environment [18]

- *Malaria*: Average malaria transmission stability index value in the  $0.25 \times 0.25$  degree cell. Data from [Kiszewski et al. \(2004\)](#).
- *Tsetse*: Average tsetse fly suitability index value in the  $0.25 \times 0.25$  degree cell. Data from [Alsan \(2015\)](#).

## Appendix B. Additional Figures

### B1. Annual Numbers of Witwatersrand Mine Workers from Southern Mozambique

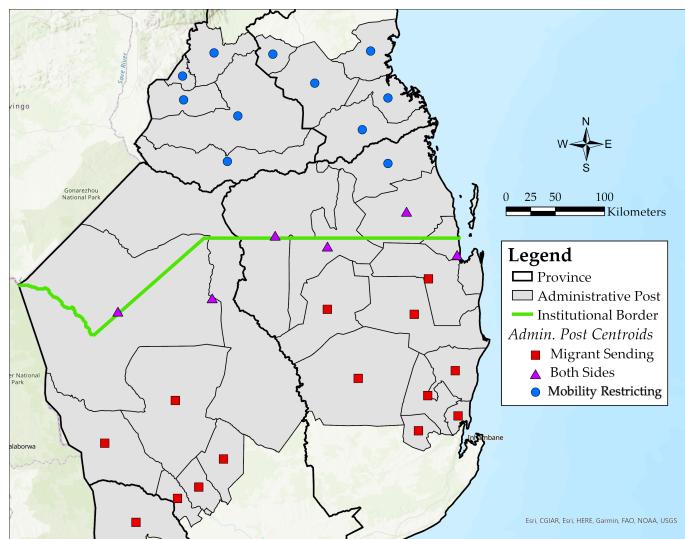
**Figure B1:** Southern Mozambican Men Received by Mines (000s), 1920-89 [9, 11]



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 mobility-restricting institution, and the black line in 1975 denotes Mozambique's independence from Portugal and deterioration of relations with South Africa.

### B2. Map of Administrative Posts

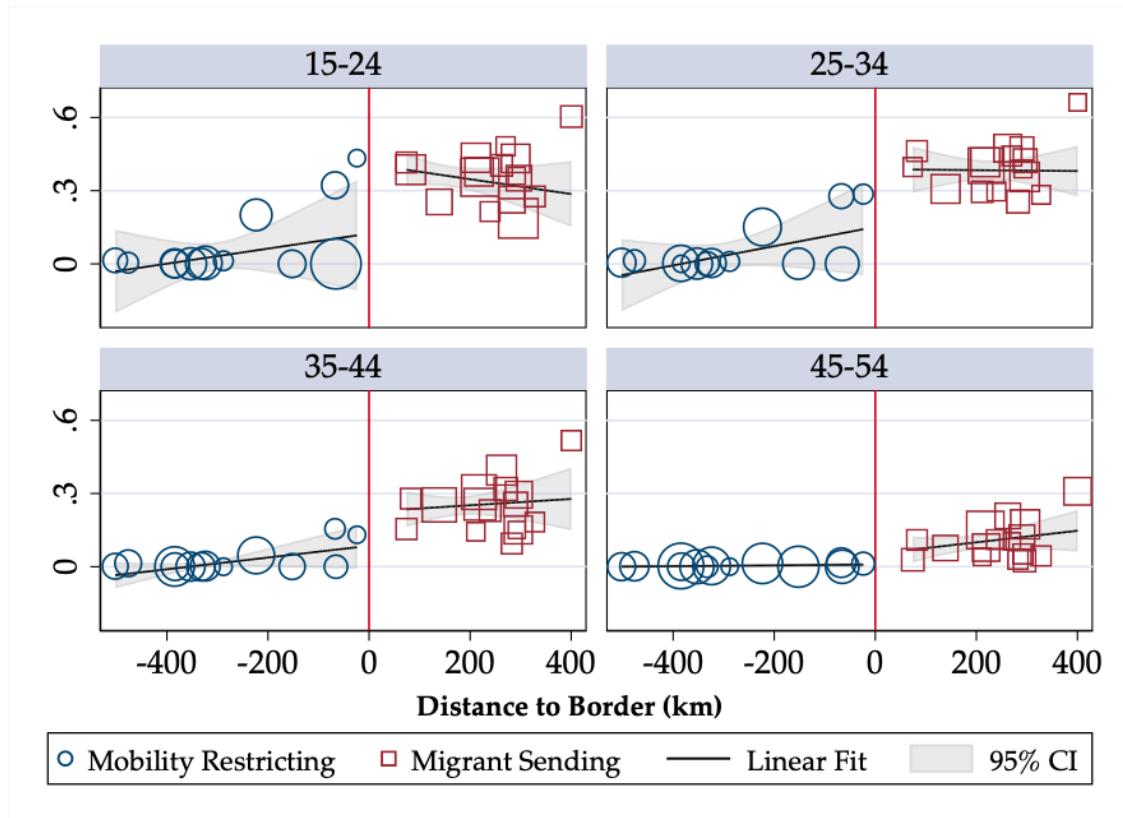
**Figure B2:** Map of Administrative Posts [17, 31]



Notes: Map shows administrative posts with centroids within 200 km of the border. The underlying shapefile is from [Minnesota Population Center \(2020\)](#).

### B3. Heterogeneity in Circular Migration by Age

**Figure B3:** Men's Circular Migration Rates by Age Group [19]



## Appendix C. Additional Tables

### C1. HIV Blood Test Refusals

**Table C1:** HIV Blood Test Refusals [27]

	Refused Blood Test		
	Pooled (1)	Women (2)	Men (3)
Migrant Sending	0.009 (0.006) [0.006]	0.010 (0.006) [0.006]	-0.006 (0.006) [0.007]
Observations	478	500	141
Clusters	13	20	12
Bandwidth	148.4	176.8	132.8
Wild Cluster Bootstrap $p$	0.428	0.195	0.432
Spatial Autocorrelation	0.31	0.28	0.67
Spatial Autocorrelation SD	0.27	0.21	0.24
Mobility Restricting Mean	0.009	0.004	0.000

*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](#)).

## C2. Ruling Out Other HIV Risk Factors

**Table C2:** Ruling Out Other HIV Risk Factors [34]

	Genital Ulcer in Last Year		Polygynous Partnership		Forced Sex		Paid for Sex		Health Decider		Medical Circumc.	
	Women (1)	Men (2)	Women (3)	Men (4)	Women (5)	Men (6)	Women (7)	Men (8)	Women (7)	Men (8)	Women (7)	Men (8)
Migrant Sending	0.061 (0.028)	0.027 (0.014)	-0.074 (0.065)	0.017 (0.049)	-0.024 (0.043)	0.015 (0.081)	-0.057 (0.094)	0.031 (0.175)				
Observations	414	235	441	202	220	196	332	443				
Clusters	19	26	28	42	21	38	22	54				
Bandwidth	54.7	85.0	92.3	130.8	90.7	148.2	92.8	152.8				
Wild Cluster Bootstrap <i>p</i>	0.077	0.216	0.506	0.733	0.630	0.838	0.650	0.846				
Spatial Autocorrelation	-0.17	-0.13	-0.22	-0.38	-0.25	0.21	0.06	-0.15				
Spatial Autocorrelation SD	0.16	0.15	0.16	0.11	0.16	0.11	0.18	0.09				
Mobility Restricting Mean	0.005	0.008	0.337	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](#)).

## Appendix D. Excluding DHS Clusters within 2 km of the Border

As discussed in Section 5.2, I exclude the DHS survey clusters within 2 km of the border as they may have been displaced across it. In addition, because these clusters are located beachside resort city of Vilankulo, including them in the estimation may lead to significant distortion. This city of 20,000 people has been popular with international tourists since shortly after the end of Mozambique's civil war in 1992 ([Mozambique News Agency, 1999](#)). It also has had multimillion-dollar infrastructure upgrades in the last decade, including the construction of an international airport that can handle 200,000 passengers per year ([Mozambique News Agency, 2000, 2011](#)).

While this recent history is unrelated to the differences between the institutions, it may affect outcomes of interest for clusters located there. First, the presence of a tourism industry could change the incentives to invest in human capital relative to other areas in the study. It also could attract workers from other areas with high human capital, and it could lead to higher wages to those living there even if they are not involved in tourism. Finally, the new infrastructure could enable additional commerce. Therefore, its inclusion would likely distort the RD estimation for outcomes related to economic development, marriage markets, and HIV.