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

Jon Denton-Schneider†

University of Michigan

November 22, 2021

Most Recent Version Here

Abstract

During the colonial period, several types of extractive institutions organized much of economic activity in Sub-Saharan Africa. Historians argued that two of them—one pushing men into circular migration and one restricting their mobility—had different effects on marriage markets with bride price but not on development. Specifically, young men in a migrant-sending institution could make bridewealth payments, narrowing the spousal age gaps that would later increase HIV risk. But much of their wages was captured by elders through inflated bride prices instead of being saved, limiting circular migration's development potential. To compare these extractive regimes, I exploit the arbitrary border within Mozambique that separated the two institutions for a half-century (1893-1942). In the colonial era, spousal age gaps were smaller in the migrant-sending region, even after the border was erased and circular migration rates converged. Today, HIV prevalence is substantially lower in this area, likely due to narrower age gaps between partners, but development outcomes are similar. These results show how different forms of colonial extraction affect health and wealth in Africa, and that marriage markets are a channel through which historical events shape the present.

Keywords: Extractive Institutions, Circular Migration, Bride Price, Age Gap, HIV

JEL Classification: D02, I15, J12, O15, N47

*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, Price Fishback, Omar Galárraga, Ezra Goldstein, Elisa Maffioli, Jennifer Mayo, Alexander Persaud, Tauhid Rahman, Ana Reynoso, Elyce Rotella, Augustin Tapsoba, Laura Taylor, and conference and seminar participants at Arizona, Michigan, NEUDC 2021, the 2021 NBER DAE Summer Institute, DemSemX, and the 2019 Economic History Association Annual Meeting 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.

†University of Michigan, Department of Economics, 238 Lorch Hall, 611 Tappan Ave, Ann Arbor, MI 48109. Email: jdds@umich.edu. Website: jondentonschneider.com.

1. Introduction

After European powers partitioned Africa at the end of the nineteenth century, they began extracting wealth from their colonies by imposing several types of institutions. Amin (1972) underscored the significance of these regimes in African history and development by grouping present-day countries into "macro-regions" where one predominated.[1] 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 settlers' farms so that the colonial state could heavily tax their wages; and 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, creating a captive pool of low-wage labor (Roberts, 2017, p. 585).[2] Figure 1 highlights these groups of countries.

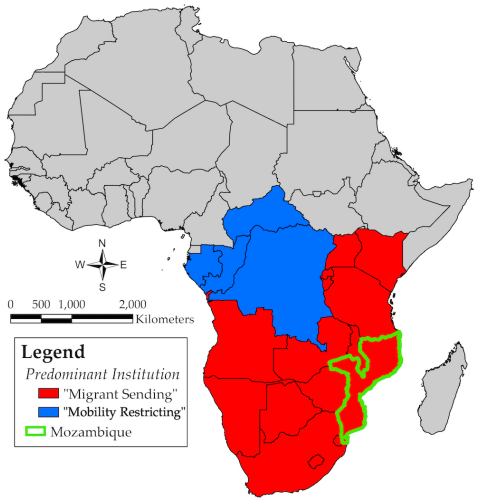
Because the two institutions (henceforth migrant-sending and mobility-restricting) organized much of colonial economic activity in the respective macro-regions, they likely had significant impacts on the historical trajectories of these parts of Sub-Saharan Africa. It is notable that today, the countries there have many of the world's highest HIV prevalence rates and are among its poorest. To effectively combat these global health and wealth disparities, policies must be suited to work in and address the social and historical contexts giving rise to them (Nunn, 2020). Understanding whether major elements of colonial history like these extractive regimes played a role—and if so, through which channels—is an important step in that direction.

However, there is no causal evidence on the impacts of the decision to establish one form of extraction instead of another in a colony. Throughout what is now the developing world, the choice was almost never between establishing an extractive institution

[1] Importantly, each macro-region contained more than one major colonizer, and each major colonizer was represented in more than one macro-region.

[2] 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 other two, I focus in this paper on the sharper distinctions between these regimes.

Figure 1: Institutional Macro-Regions in Africa



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

or one promoting inclusive prosperity (in the Acemoglu and Robinson, 2012, terminology). Instead, it was between *different forms of extraction*. Therefore, comparing these regimes and tracing their impacts through history can shed light on why disparities in HIV prevalence and poverty exist between and even within African countries.

The main challenge in generating such evidence is that institutions were not randomly assigned (Acemoglu, Johnson and Robinson, 2001). Instead, it was most likely distinct human and natural resource geographies in the macro-regions that led their colonizers to rely primarily on one type of extractive regime. This relationship complicates simple comparisons between them, and such factors (e.g., climate, crop suitability) are also likely to have directly affected many outcomes of interest.

The point of departure for this paper is that while a regime may have predominated in a macro-region, there were unique colonies like Mozambique in which it was not the only kind imposed (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 them was arbitrary, consisting almost entirely of straight lines defined by latitude and longitude.

In this paper, I exploit the arbitrary border between one of Africa's most important migrant-sending institutions (1897-1965), which pushed 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). I provide greater detail regarding these institutions and the differences between them in Section 2, and Figure 2 shows them on a map of Mozambique. According to historical accounts, their most important difference was indeed in men's labor mobility, which was high in the migrant-sending region but restricted across the border to create a pool of low-wage, conscriptable labor.

Because the ethnic group split by the institutional border practiced bride price, historians also noted that the higher wages from circular migration allowed young men to marry earlier than was previously possible. However, they argued that two factors limited the impact of circular migration on economic development. The massive increase in the number of young men who could enter the marriage market led to higher bride prices paid to in-laws, consuming much of their wages (Harries, 1994). In addition, parents faced incentives to use daughters' bride prices to acquire wives for unmarried sons, whose in-laws then faced the same incentives (Junod, 1912).[3] The result was circular migrants' earnings in large part being locked into a cycle of bridewealth transfers between elders rather than being saved by young men.

Therefore, the narrative evidence suggests the institutions led to lasting differences in marriage markets but only limited differences in development outcomes. The first empirical portion of the paper tests this hypothesis in the colonial era. As I describe in Section 3, I digitized detailed data from district-level summaries of two censuses and georeferenced them to maps of colonial Mozambique. I use these data and exploit the arbitrary border between the two institutions in a geographic regression discontinuity (RD) design, which estimates the causal impact of a region's historical assignment to the migrant-sending institution relative to the mobility-restricting one.

[3] Specifically, sons' wives cultivated additional land and populated it with children. Also, because sons who relied on elders for bridewealth became their cadets, it expanded elders' tax bases (Harries, 1983).

In Section 4, I study the outcomes that were different along the border two years before 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: working-age men in the migrant-sending institution were 20 percentage points (p.p.) more likely to engage in circular migration, which is a very large increase relative to the average rate of 5 p.p. in the mobility-restricting region. However, circular migration rates had converged by 1960, as labor mobility for men became much easier after the institutional border was erased.

In spite of this convergence, marriage market outcomes remained markedly different nearly two decades after the end of the mobility-restricting institution. In 1940, the ratio of married men to married women in young adult age groups was 30 p.p. higher just inside the migrant-sending region, though the difference shrank to 10 p.p. by 1960. These results suggest that the length of time circular migration had been possible for was an important determinant of marriage market outcomes.

To compare (correlates of) economic development in this period, I also examine school enrollment rates given the links between migration and human capital investment (e.g., Yang, 2008; Dinkelman and Mariotti, 2016).[4] Though the results are difficult to interpret given the history of schooling provision in the two institutions, I find slightly lower rates of schooling for boys in the migrant-sending area in 1940.[5] But two decades later—once circular migration rates and the provision of schooling were the same across the border—enrollment rates had converged, consistent with the absence of lasting differences in this aspect of economic development.

To rationalize the slow equalization of marriage outcomes and to conceptualize how it could affect the present day, I develop in Section 5 an overlapping generations model

[4] No direct measures of wealth were collected in either of the colonial censuses, likely because it would have been very difficult to do so accurately. In contrast, enrollment was easily measured.

[5] Schooling for Africans was generally provided by Protestant missionaries in the migrant-sending region and Catholics in the mobility-restricting one, though it was unavailable to the overwhelming majority of Black children under either institution. After the mobility-restricting regime ended, Catholic missions began to provide it in both regions (Morier-Genoud, 2019; Fernández Cebrián, 2021).

of a marriage market with bride price that adapts the Tertilt (2005) framework. To match the context under study, wages depend on age and sex, fathers transfer their daughters' bride prices to her brothers so they can marry in the next period, and older generations are smaller than younger ones due to mortality. At baseline, young men earn less than young women, who in turn earn less than old men. The bride price is thus greater than young women's wages (Corno, Hildebrandt and Voena, 2020), making it unaffordable for young men.[6] As a result, all marriages are between old men and young women.

I then study what happens to the marriage market after a shock that raises wages for a share of young men well above those of old men (i.e., circular migration).[7] In the first period, there is a large increase in demand for brides but the supply of young women is fixed. Therefore, the share of marriages that are cross-generational falls. This pattern is reinforced in the second period by two factors: men who married last period do not marry again in old age, and the young generation is larger than previous ones because of incentives to have more children.[8] Only by the third period can the market reach its new steady state, which may be 90 to 105 years later if a period is 30 to 35 years.

As the shock to men's wages in the mobility-restricting region only began in 1942, this slow transition could mean that there are fewer age-disparate relationships in the migrant-sending institution. Because age gaps between partners are an important risk factor for HIV's spread (de Oliveira et al., 2017; Schaefer et al., 2017), the implication is that the virus's prevalence is lower on this side of the former border.[9] Conversely, due to the convergence in circular migration, this HIV risk factor should be the same along the border (Weine and Kashuba, 2012); the same is true for schooling. These equalizations also imply that there should not be differences in these aspects of economic development.

[6] Only young women are fecund in this model (e.g., Siow, 1998). Because the purpose of marriage is procreation, all young women are married off before they lose value on the marriage market (Tertilt, 2005).

[7] Old men rarely were circular migrants because mine labor required one to be in his physical prime.

[8] The former is due to the disutility from work in old adulthood—they can consume from their wives' wages—and the larger bridewealth transfers to his unmarried brothers price him out of the market.

[9] Age gaps raise transmission because older men have been sexually active for longer, so they are more likely to have been infected. Relationships with young women spreads HIV into the next generation.

As I show in Section 6, the present-day results are consistent with my 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. This large effect is remains constant even when splitting the sample by sex. I also examine age profiles of HIV prevalence by sex near the border and show that they are consistent with the effect size. Additionally, as I hypothesized, I find no differences at the border today in wealth, schooling, or children's health, suggesting equal levels of development.

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.

As such, this paper contributes to four main literatures. First, a number of influential studies examine the effects of colonial institutions on modern outcomes (Acemoglu, Johnson and Robinson, 2001; Banerjee and Iyer, 2005; Dell, 2010; Michalopoulos and Papaioannou, 2014, 2016), and there is recent evidence specifically on concessions (Dell and Olken, 2020; Lowes and Montero, 2021*a*; 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 a century 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; Khanna, Theoharides and Yang, 2020). I contribute specifically to the literature on migration's effects on human capital, especially health (Black et al., 2015; Dinkelman and Mariotti, 2016). 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 add to our understanding by documenting the long-run consequences of short-term labor mobility for young African men. I also show which outcomes converge and which ones remain different after the ability to engage in circular migration had equalized.

In addition, the economic analysis of non-Western marriage markets—which determine how most of the world marries—is an expanding area of study (Tertilt, 2005; Corno, Hildebrandt and Voena, 2020; Reynoso, 2021). I contribute in this field by showing how they interact with labor market conditions to shape behaviors affecting human capital in the long run (Chiappori, Iyigun and Weiss, 2009; Greenwood, Guner and Vandenbroucke, 2017; Ashraf et al., 2020), though in my case I focus on its health component. I also show how marriage markets with asset transfers, and family life in developing countries more broadly, have important effects on health and longevity (Calvi, 2020). Thus, they are vital to understand when crafting policies to remedy these consequences.

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, 2021*b*). Because HIV's spread 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 (Iliffe, 2006; 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, organizing the analysis with a theoretical model, and using the historical record to make a case for this 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 explosion of HIV in the late 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 sixteenth-century port cities 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.[10] 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).[11] 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 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

[10] 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.

[11] 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).

> 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).[12] 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).[13] To keep wages low, the mining companies formed the monopsonistic Witwatersrand Native Labour Association (WNLA) to recruit workers on their behalf.

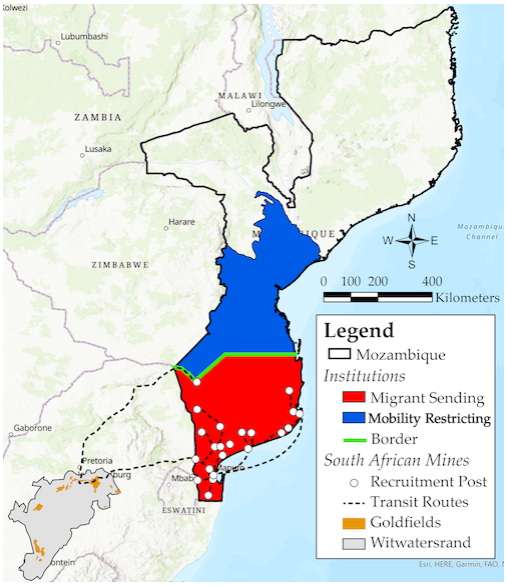
WNLA and Portuguese authorities signed several agreements beginning in 1897 that formalized recruitment in the colonial state's southern territory. The government derived revenues from all parts of this process: licensing fees from recruiters, payments from WNLA for each worker, and permit fees from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

These agreements also regulated miners' contracts and how they were paid. Contract

[12] 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.

[13] 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).

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).[14] The Portuguese had long argued for this provision because miners spent much of their wages on the Witwatersrand—often to buy status goods—rather than in Mozambique (Harries, 1994).

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

[14] 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).

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 pool of low-wage 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, 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." 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."[15]

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

[15] One 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). 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). 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) 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.

(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.

Some historians have argued that the bridewealth system limited the economic impact of circular migration in southern Mozambique. Because elders controlled social life, especially marriage, migrants' wages may have been "encapsulated within the sphere of circulating bridewealth controlled by [them]. As bridewealth was kept in trust . . . to provide future generations with the means of acquiring wives, . . . it could not be invested," reducing much of its potential effects on development (Harries, 1983, p. 321).

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.[16]

In spite of this institution's end, the extraction of wealth from labor continued throughout the colonly until the end of Portuguese rule in 1975 (Isaacman et al., 1980; Guthrie, 2016). Nonetheless, the newly-independent Mozambique 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 in this period, and shortly after stability returned its HIV epidemic began to explode (Audet et al., 2010).[17]

[16] The Salazar regime also unified education policy across Mozambique at this time, having the Catholic Church take over—and greatly increase—schooling for Africans (Fernández Cebrián, 2021).

[17] 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

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.

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 Mozambican War

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).

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 development as well as the ease of measuring it accurately.

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.[18] 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").[19] 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.[20]

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.[21]

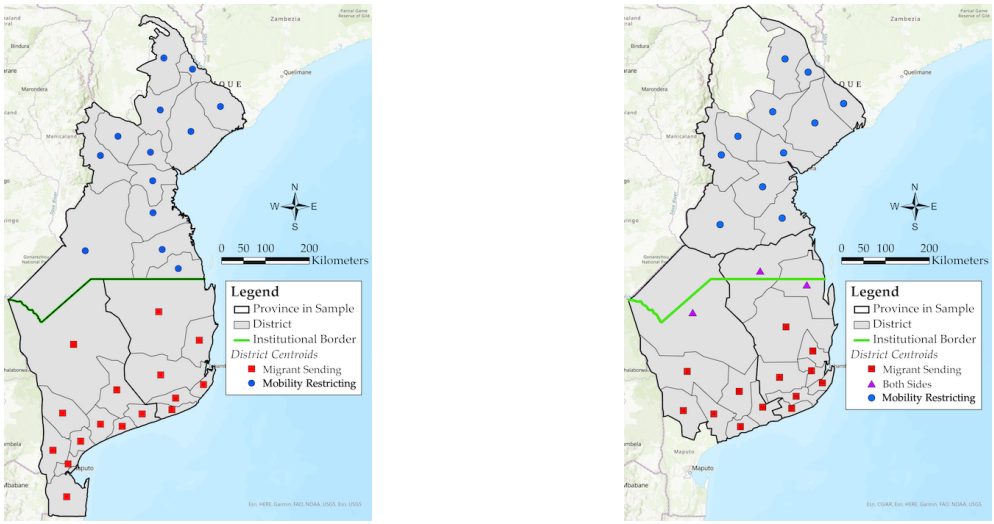
[18] Each census's questionnaire asked whether a man worked abroad but the 1960 summary tables grouped circular migrants into a category with all men who worked in a mine regardless of its location. However, nearly all men in this category worked abroad, so I consider it a measure of circular migration.

[19] The second ratio approximates the number of children born to women of reproductive age.

[20] The numerator excludes those who had left school before enumeration.

[21] These two cities are Lourenço Marques and Beira, the capital of the restricted mobility region.

Figure 4: Maps of Georeferenced Colonial Census Data



(a) 1940 **(b)** 1960

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

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:

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

where y_d is the outcome of interest for district d in the set B defined by the bandwidth restrictions above.[22] The explanatory variables are MigrantSending_d, an indicator for whether d's centroid is in that institution; $f(\text{Distance}_d)$, the RD polynomial controlling for smooth functions of a centroid's distance to the institutional border; and Lon_d, a centroid's longitude coordinate, which Kelly (2020) recommends including in RD designs to

[22] 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.

capture east-west trends.[23] 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 τ 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, Portuguese colonial officials quasi-randomly allocated the territory around it to one of the two institutions. I examine the arbitrariness of the border in the next section.

3.2.1. Addressing Concerns with Estimation and Inference

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

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

I take two steps to address this potential problem. First, I calculate 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

[23] Distance$_d$ has a near-perfect correlation with latitude ($\rho > 0.99$), so it accounts for north-south trends.

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.[24] To generate the spatial weighting matrix, I set the bandwidth so that each district has at least one other neighbor, impose a Bartlett kernel, and row standardize the entries so that I is between -1 and 1.[25] I report the difference of the observed and expected I-statistics as the measure of spatial autocorrelation along with I's standard deviation.[26]

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 km $\times$ 25 km in the study area—and estimate equation (1) clustering standard errors by third-level administrative unit ("administrative posts," shown in Appendix B2).[27] 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.

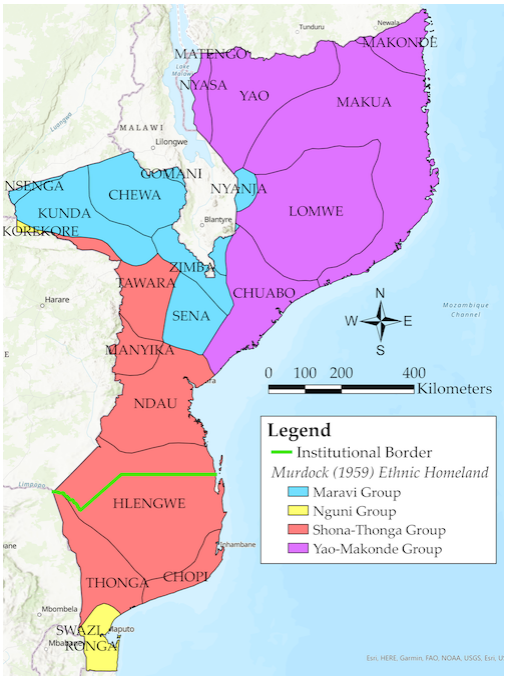
[24] The I-statistic's expected value under no spatial autocorrelation is $\frac{-1}{N_s-1}$, which approaches zero from below as the number of unique sites (N_s) increases. An observed I greater (less) than the expectation indicates positive (negative) spatial autocorrelation, meaning neighboring sites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

[25] The resulting bandwidth is approximately 170 km for 1940 and approximately 140 km for 1960.

[26] The asymptotic distribution of the I-statistic is standard normal.

[27] See Section 6.2 for details on RD bandwidth selection when using geographically disaggregated data.

Figure 5: Ethnic Group Homelands in Mozambique



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

Table 1: Balance Tests at the Border

	Geographic Traits				*Disease Suitability*	
	Elevation (1)	Rainfall (2)	Slope (3)	Soil Index (4)	Malaria (5)	TseTse (6)
Migrant Sending	-1.798	1.153	0.014	3.207	-0.270	-0.002
	(31.917)	(6.533)	(0.100)	(3.887)	(0.368)	(0.009)
	[23.254]	[4.045]	[0.054]	[2.658]	[0.244]	[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 × 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.

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 end of the mobility-restricting institution (Panel A) and 18 years after (Panel B), and Figure 6 presents selected RD plots. The results show that the institutions differed substantially in circular migration and marriage while they both existed, but only the latter remained markedly different nearly two decades after the mobility-restricting regime ended.

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 were any men circular migrants.

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 men worked abroad.[28] This pattern implies that the regime significantly constrained men's mobility and its end in 1942 led to major changes in their occupational choices.

[28] I exclude a migrant-sending district with only one circular migrant, which is most likely an error.

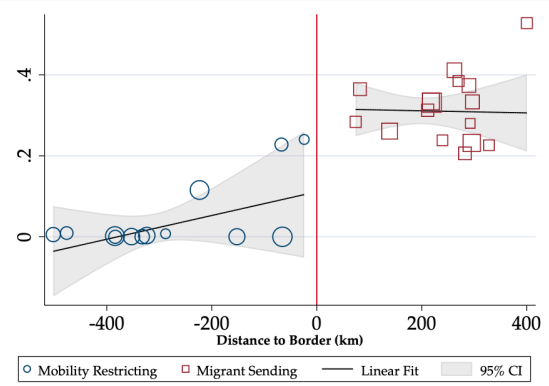
Table 2: Comparing Institutions during the Colonial Era

	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	0.015	0.269	0.312	0.202	-0.028	-0.001
	(0.089)	(0.018)	(0.080)	(0.113)	(0.098)	(0.016)	(0.006)
	[0.085]	[0.016]	[0.061]	[0.088]	[0.092]	[0.015]	[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	0.006	0.102	0.106	0.056	-0.014	-0.026
	(0.049)	(0.004)	(0.057)	(0.147)	(0.086)	(0.033)	(0.024)
	[0.050]	[0.004]	[0.053]	[0.131]	[0.082]	[0.029]	[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

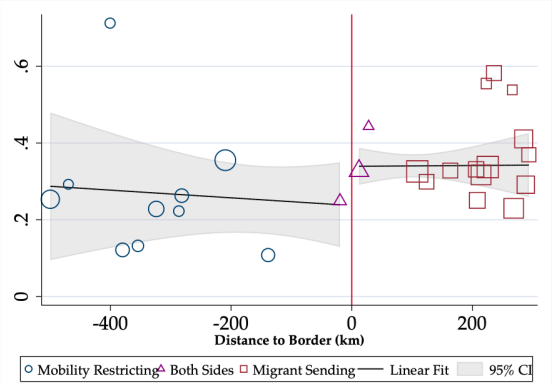
Notes: Observations are districts. Robust standard errors are in parentheses and Conley standard errors using a 100-km bandwidth and a Bartlett kernel are in brackets. Regressions estimate a local linear RD specification on each side of the border using a triangular kernel and include longitude as a 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.

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 predated the institutions and enabled men's circular migration in the first place.

Figure 6: RD Plots for Colonial-Era Differences between Institutions



(a) Men Circular Migrants, 1940



(b) Marriage Ratio Ages 15-24, 1960

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. Data sources, variable definitions, and RD bandwidth selection criteria are in Section 3.

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 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 existed, the marriage ratio for both age groups increased by around 0.3 just inside the migrant-sending region.[29] 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.

[29] 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.

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.[30] The positive response of fertility to circular migration (and its higher wages) suggests that Malthusian dynamics were at work, which is 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

[30] 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.

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. Linking Past and Present

In this section, I develop a simple overlapping generations model of a marriage market with bride price and polygyny in a Malthusian setting to make predictions about how a source of high wages for young men (i.e., circular migration) affects marriage, fertility, and living standards. The goal is to rationalize why marriage market outcomes would take longer to converge than circular migration rates, and why we might expect marriage differences to continue to today but not necessarily those in standards of living.

To do so, I draw on the models of Sub-Saharan African marriage markets in Tertilt (2005) and Corno, Hildebrandt and Voena (2020) as well as the Malthusian economies in Galor and Weil (2000) and Ashraf and Galor (2011). To match the specific context and convey the intuition in the simplest possible manner, I abstract away from the role of capital in subsistence agricultural production, eliminating the savings decision that plays a large role in the Tertilt (2005) model.[31]

5.1. Setup

The model takes place over infinite discrete time. Individuals are either male (M) or female (F). Men make all decisions in the model: for himself if single or for his household (himself and each wife and unmarried daughter) if married.[32] After individuals

[31] Tertilt (2005) studies the macroeconomic consequences of polygyny and simulates the impact of a ban on the practice, which is beyond the scope of this paper. Also, because the male side of production—the Witwatersrand mining companies or domestic employers using forced labor—received investment from European capital markets rather than agents in the model and women worked in subsistence agriculture, it is reasonable that Black Mozambicans' savings would have little to no role in this context.

[32] In the case of widows, I assume they run their late husbands' households exactly as they would have.

are born, they live for up to 2 periods of adulthood, young (Y) and old (O). Surviving childhood is determined by parents' choices (discussed below), whereas the probability of young adults surviving into old adulthood is fixed at $1-\mu$. Men are always fecund while only young women are fecund.[33] The discount factor is $\beta \in (0,1)$.

5.1.1. Timing, Decisions, and Cultural Practices

At the beginning of a period, young men must work due to forced labor laws while old men can choose whether to do so, incurring a disutility δ if they do.[34] Those who work earn age-dependent incomes y_t^A, $A \in \{Y,O\}$, which are net of taxes and any costs incurred to earn them (e.g., migration).

After men work, they and (the parents of) single women enter a frictionless marriage market. Men demand $w_t^A \geq 0$ wives and parents supply $d_t^A \geq 0$ of their daughters as brides, giving their husbands control over their reproductive capacity and labor income.[35] At the end of the period, men pay bride prices p_t to their in-laws.[36]

Her parents must then transfer all of their daughters' bride prices to the chief whose authority they are under. He holds these amounts in trust and divides them equally among her surviving unmarried brothers at the beginning of the next period. I do not model how this practice arose but simply impose it on the marriage market to match the historical context discussed in Section 2.

After the marriage market closes, women work in subsistence agriculture and produce y_t^F. Married men with fecund wives then choose to have $2n_t^A$ children with an equal sex ratio. As in Tertilt (2005), polygynous men split their children equally across his wives, and they survive at a total cost of $\frac{\psi(n_t^A)^2}{w_t^A}$.[37] After they are born, households consume, all old adults and the fixed share of young adults die, and the period ends.

[33] This asymmetry makes fecund women scarce (Siow, 1998)
[34] This term could reflect the greater discomfort from arduous physical labor in old age.
[35] While a wife or a daughter is clearly indivisible, these quantities are best thought of as the average number of wives demanded or brides supplied by men in a generation.
[36] In Southern Africa, bride prices can be paid after a marriage occurs (Ansell, 2001).
[37] Note that this cost increases in the number of children and decreases with the number of wives.

5.1.2. Preferences and Budget Constraints

Men value their own or their household's consumption in each period with log preferences. In this patrilineal society, they also gain utility from continuing their lineages by having sons who marry. While he cannot make decisions for his sons, a father influences whether they marry by marrying off daughters, ensuring that his sons (i.e., their brothers) will have more resources available when they enter the marriage market in the next period. This decision is incorporated into men's utility functions as $\lambda \log(d_t^A)$, $d_t^A \in [0, n_t^A]$, where λ captures how important lineages are.[38]

It is immediately clear that men will marry off all of their daughters ($d_t^A = n_t^A$), which makes the last term in the utility function a preference for fertility. Therefore, fathers will marry all daughters off while they are young women because potential sons-in-law will not pay for infecund wives.

If incomes for young men are low enough, they will not be able to marry when young and must incur the discomfort of working in old age to marry then.[39] As a result, they choose to consume their entire incomes when young. Substituting in the period budget constraint, their problem after surviving to old age in period t is

$$\max_{w_t^O, n_t^O \geq 0} \log\left[y_t^O + \frac{p_{t-1}}{1-\mu} - w_t^O(p_t - y_t^F) - \frac{\psi(n_t^O)^2}{w_t^O}\right] + \lambda \log(n_t^O), \tag{2}$$

where $\frac{p_{t-1}}{1-\mu}$ is his share of his sisters' bride prices from their marriages in $t-1$.

Conversely, if young-age incomes are high enough for some men and the pain from working in old age is large enough, they will marry only when young and choose to consume from their wives' production when old.[40] Their problem as young men in t is

[38] This preference mirrors the one for children in Tertilt (2005) and Ashraf and Galor (2011). While old men die before they can marry off new daughters, the widows who will run their households will behave exactly as they do. Thus, it is effectively equivalent and more convenient to write the problem this way.

[39] A formal statement of this condition is in the next section.

[40] Formally, it requires that this pain outweighs the gains from more consumption and children: $\delta > \log\left[y_{t+1}^O + w_t^Y y_{t+1}^F - w_{t+1}^O(p_{t+1} - y_{t+1}^F) - \frac{\psi(n_{t+1}^O)^2}{w_{t+1}^O}\right] - \log(w_t^Y y_{t+1}^F) + \lambda[\log(n_t^Y + n_{t+1}^O) - \log(n_t^Y)]$.

$$\max_{w_t^Y, n_t^Y \geq 0} \log\left(y_t^Y - w_t^Y(p_t - y_t^F) - \frac{\psi(n_t^Y)^2}{w_t^Y}\right) + \beta(1-\mu)\left[\log\left(w_t^Y y_{t+1}^F\right) + \lambda \log(n_t^Y)\right]. \quad (3)$$

Because he married when young, he does not receive any of his sisters' bride prices in old age—it gets redistributed to his unmarried brothers. Their problem as old men is similar to (2) except their share gets scaled up by the fraction who did not marry.

5.1.3. *Subsistence Agricultural Production*

Women produce subsistence agricultural output using a constant-returns-to-scale technology with labor and land as inputs, $Y_t^F = (AX)^\alpha [N_t + (1-\mu)N_{t-1}]^{1-\alpha}$, $\alpha \in (0,1)$, where A is the level of technology, X is land, and $N_t + (1-\mu)N_{t-1}$ is the number of young and old women.[41] As in Galor and Weil (2000), there are no property rights over land, so its return is zero and women's wages are the average product

$$y_t^F = \left(\frac{AX}{N_t + (1-\mu)N_{t-1}}\right)^\alpha \equiv x_t^\alpha. \quad (4)$$

5.1.4. *Population Dynamics and Marriage Market Clearing*

The number of young adults of each sex in a period is determined by the reproductive choices of men married to young women in the previous period. Its law of motion is

$$N_{t+1} = N_t n_t^Y + (1-\mu)N_{t-1} n_t^O. \quad (5)$$

Along similar lines due to the preferences discussed in the previous section is the marriage market clearing condition. Specifically, because fathers will marry off every daughter when she is young, equating the supply of brides with demand for them yields

$$N_t = N_t w_t^Y + (1-\mu)N_{t-1} w_t^O. \quad (6)$$

[41] As Ashraf and Galor (2011) describe, A reflects the soil, climate, and practices used in production.

5.2. *Predictions*

I solve for the equilibria of this marriage market and economy defined by the optimal choices in (2) and (3) taking the female wage in (4) and the bride price p_t as given, population evolving according to (5), and the marriage market clearing as in (6). My interest is in the steady state under low wages for young men, and what happens after high wages become available to a subset of young men.

5.2.1. Initial Steady State: Low Wages for Young Men

Suppose first young men's wages are fixed at a level y_0^Y that is less than young women's, which in turn are less than old men's: $y_0^Y < x_t^\alpha < y_0^O$.[42] The first-order conditions for (2) yield period-t old men's choices

$$n_t^O = w_t^O \frac{\sqrt{p_t - x_t^\alpha}}{\sqrt{\psi}}, \; w_t^O = \frac{\lambda\left(y_0^O + \frac{p_{t-1}}{1-\mu}\right)}{2(1+\lambda)(p_t - x_t^\alpha)}. \tag{7}$$

Substituting the expressions in (7) into (5) and (6) and solving for steady-state values denoted by 0 subscripts implies that the number of children equals the number of wives and it gives the bride price as a function of the exogenous variables:

$$w_0^O = \frac{1}{1-\mu} = n_0^O \implies p_0 = \left(\frac{AX}{N_0(2-\mu)}\right)^\alpha + \psi. \tag{8}$$

Intuitively, the number of wives married and children born must be equal for the population to be constant. Because all marriages are age-disparate, the rate of polygyny is completely determined by the level of mortality. In addition, consistent with Corno, Hildebrandt and Voena (2020), the bride price exceeds a young woman's contribution to her parents' budget in the absence of direct utility from marriage, keeping young men from entering the marriage market given their wages. Also note that a larger steady-state population implies a lower bride price because it reduces women's wages.

[42] This wage pattern could reflect returns to experience in forced labor (e.g., an overseer position).

Table 3: Summary of Marriage Predictions

Period	Age	M Marrying	F Marrying	Share w. Age Gap
0	Y		N_0	1
	O	$(1-\mu)N_0$		
1	Y	ϵN_0	N_0	$\frac{1-\mu}{1-\mu+\epsilon}$
	O	$(1-\mu)N_0$		
2	Y	ϵN_1	N_1	$\frac{(1-\mu)(1-\epsilon)}{(1-\mu)(1-\epsilon)+\epsilon\frac{N_1}{N_0}}$
	O	$(1-\mu)(1-\epsilon)N_0$		
3	Y	ϵN_2	N_2	$\frac{(1-\mu)(1-\epsilon)}{(1-\mu)(1-\epsilon)+\epsilon\frac{N_2}{N_1}}$
	O	$(1-\mu)(1-\epsilon)N_1$		

Notes: Table presents a summary of the marriage market model's predictions regarding age-disparate marriages. Period 0 denotes the baseline steady state and Period 1 is the first period in which circular migration is possible. The respective ages denote the young and old generations in a period. The sex-specific columns show the number of each sex and generation who marry in a period. The final column shows the share of women in age-disparate marriages.

5.2.2. First Periods of High Wages for a Subset of Young Men

Now assume that much higher wages y_1^Y are available to a share $\epsilon \in (0,1)$ of young men: $y_0^Y < x_t^\alpha < y_0^O < y_1^Y$.[43] In this new regime, the expressions for period-t old men's choices are the same as in (7). However, now the ϵ share of period-t young men with high wages make decisions defined by (3):

$$n_t^Y = \frac{\sqrt{\beta(1-\mu)\lambda w_t^Y}\sqrt{y_1^Y - w_t^Y(p_t - x_t^\alpha)}}{\sqrt{\psi(\beta(1-\mu)\lambda+2)}},\ w_t^Y = \frac{\beta(1-\mu)(\lambda+2)y_1^Y}{2[\beta(1-\mu)(1+\lambda)+1](p_t - x_t^\alpha)}. \quad (9)$$

Substituting the demands for wives into the market clearing condition (6) yields an expression for the current period's bride price as a function of the previous period's, which does not have a closed-form solution.

However, because my focus is on the consequences of the wage shock for a subset of young men rather than the steady state, it is instructive to examine the first few periods after the new wage regime begins. Consider the market clearing condition in the first period following its introduction (denoted as $t = 1$). After substituting in the initial

[43] This share could be determined by migration costs that decline with idiosyncratic ability, yielding a threshold ability level above which young men engage in circular migration. Given the high disutility from physical labor in old age discussed earlier, no old man would do so.

steady-state bride price, the result is a closed-form solution for the period-1 value,

$$p_1 = x_0^{\alpha} + \epsilon \frac{\beta(1-\mu)(\lambda+2)y_1^Y}{2[\beta(1-\mu)(1+\lambda)+1]} + (1-\mu)\frac{\lambda\left(y_0^O + \frac{x_0^{\alpha}+\psi}{1-\mu}\right)}{2(1+\lambda)}. \tag{10}$$

which is clearly greater than p_0 in (8) for any reasonable cost of children's survival. Even though the new wage regime has begun, the bride price is still partially a function of variables determined in the last period—namely, the female wage, both in itself and as part of last period's bride price.

This feature along with mortality and the lag between women being born and entering the marriage market lead to a noticeably delayed adjustment. As is clear in Table 3, it takes at least another generation after higher wages become available for a subset of young men for the share of age-disparate marriages to effectively stabilize at a much lower level. Importantly, the largest part of the transition occurs in the first period and the changes get smaller in the subsequent periods.

5.3. Linking Past and Present

Given the colonial-era results and the model above, it is possible that marriage market differences have continued to today. Setting the length of a generation to be 30 years implies that marriage markets in the former mobility-restricting region would experience a large but incomplete transition in the 1940 to 1970 period. In this framework, only after 2000 period had ended would marriage outcomes complete most of their convergence with the former migrant-sending region.

If these predictions are correct, there would be clear implications for HIV prevalence and economic development today: seroprevalence should be lower just inside the migrant-sending region while there should not be substantive differences in living standards along the border.

5.3.1. *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). As such, fewer age-disparate relationships should lower HIV prevalence and give its age profile a later peak, especially for women.

Concurrent sexual partnerships, as formalized by polygyny, can also increase the risk of contracting the virus (Tanser et al., 2011). An important reason is that the probability of transmission increases with viral load, which can be very high shortly after acquiring HIV (Quinn et al., 2000). Therefore, sexual contact with multiple partners in this window raises the risk that each of them will become infected.

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, as would restricted mobility during the 1977-92 civil war (Iliffe, 2006; Audet et al., 2010).

5.3.2. *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.

6. Effects of the Institutions in the Present Day

Given the framework above, in this section I study the institutions' impacts on HIV prevalence and economic development today. I first describe the modern data and the refinements to the colonial-era RD estimation strategy that I use to study these present-day outcomes. Tables 4 and 5 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 substantive differences in measure of living standards today.

6.1. 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 for respondents' anonymity and privacy.[44]

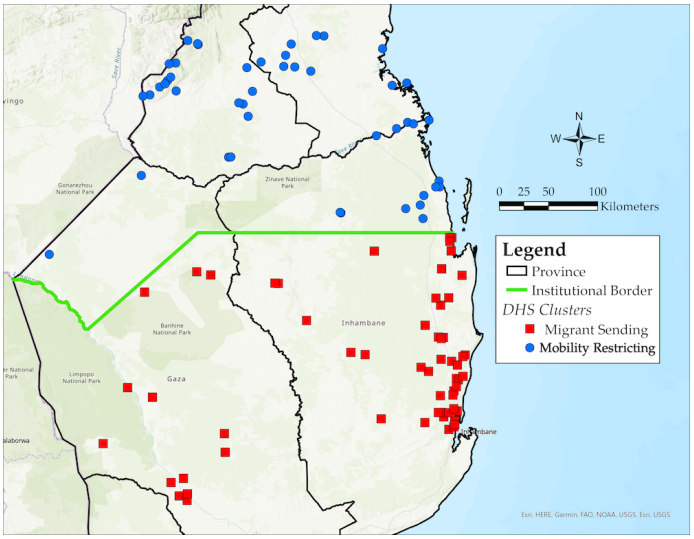
As such, it is possible that four urban clusters along the coast have been displaced into the wrong institution. For this reason and others related to the city's recent history that I discuss in Appendix D, I remove these clusters from the sample. After doing so, all of the remaining ones are in the 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.[45] For economic development, the variables of interest are an index of

[44] 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.

[45] 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).

household asset ownership (measured in 2009, 2011, 2015, and 2018), an indicator for whether a child is stunted (2011), and years of schooling (2009, 2011, and 2015).[46]

6.2. *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:

$$y_{i,c} = \alpha + \tau \text{MigrantSending}_c + f(\text{Distance}_c) + \text{Lon}_c + \mathbf{X}_i\beta + \delta_t + \epsilon_{i,c} \quad \text{for } c \in B^*_{\text{MSE}} \quad (11)$$

where $y_{i,c}$ is an outcome for individual i in DHS survey cluster c and the first three right-hand side variables are as before. I also include the vector $\mathbf{X}_i$ containing individual-level controls (age, age squared, and a female indicator) and the survey-year fixed effect δ_t.

[46] 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^*_{MSE}. I continue to use a local linear RD specification with a triangular kernel.

6.2.1. Addressing Concerns with Estimation and Inference

An estimation issue arises from the displacement of clusters mentioned earlier. Because the displacement is random, it induces classical measurement error in the running variable, biasing the RD coefficients toward zero. For inference, I cluster standard errors by DHS survey cluster. However, one concern with this approach 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 concern is spatial autocorrelation, which again I address by calculating Conley standard errors and Moran I-statistics as discussed previously.[47]

6.3. Results: HIV Prevalence

I first examine the spatial distribution of HIV among adults along the institutional border. Table 4 Column (1) pools both sexes and shows that adult HIV prevalence drops 10 p.p. just inside the migrant-sending institution. This point estimate is large relative to the 22 percent of the 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

[47] 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 4: HIV Prevalence

	Positive Blood Test		
	Pooled (1)	Women (2)	Men (3)
Migrant Sending	-0.103	-0.089	-0.097
	(0.049)	(0.053)	(0.084)
	[0.037]	[0.041]	[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).

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, 2021*b*).

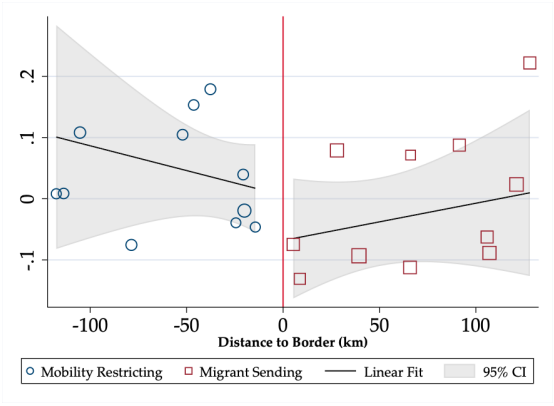
To rationalize the large effect sizes estimated above, I compare the age profiles of HIV prevalence in the two institutions.[48] Specifically, I calculate the mean seroprevalence for each 10-year age group within each sex's MSE-optimal RD bandwidth.[49] Figures 8c and 8d plot these age profiles. A clear 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) transmission cycle and its implications discussed in Section 5. The peak of women's seroprevalence in the former mobility-restricting institution is both larger than the one

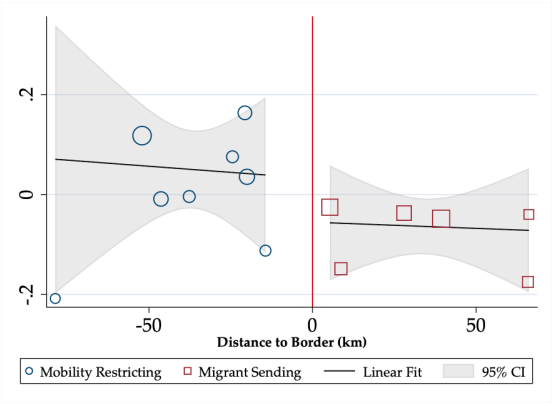
[48] Note also that prevalence is a stock, not a flow, and even small differences in transmission rates can generate large differences in the size of an epidemic (Viboud, Simonsen and Chowell, 2016).

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

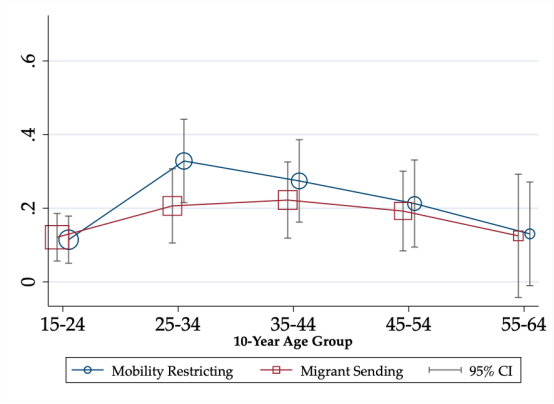
Figure 8: HIV Prevalence RD Plots and Age Profiles



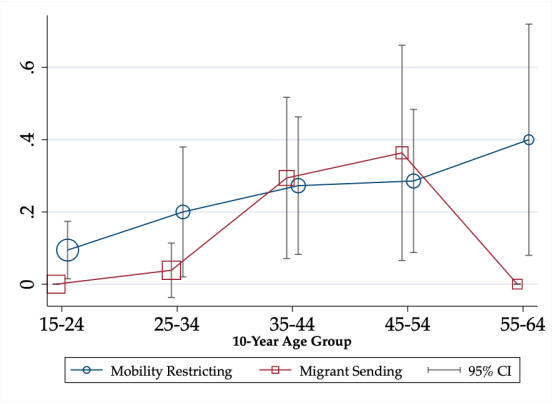
(a) RD Plot: Women



(b) RD Plot: Men



(c) Age Profile: Women



(d) Age Profile: Men

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.

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.

6.4. Results: Economic Development

Next, I compare economic development outcomes in the former institutions. The point estimates in Table 5 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 mobility-restricting means. The coefficients for childhood

Table 5: Economic Development Outcomes

	Assets	*Stunting*	*Years of Schooling*	
	Index	Children	Females	Males
	(1)	(2)	(3)	(4)
Migrant Sending	0.067	-0.055	0.377	0.224
	(0.322)	(0.117)	(0.327)	(0.782)
	[0.414]	[0.132]	[0.281]	[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).

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 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.[50] 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).[51]

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.[52] 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.[53]

Table 6 reports the results of estimating equation (11) for age disparities in these datasets and Figure 9 presents RD plots for these outcomes. Column (1) shows that the

[50] 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.

[51] 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.

[52] 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.

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

Table 6: Age-Disparate Partnerships

	Census	*DHS*	
	Spouse	*Last Sex Partner*	
	Women (1)	Women (2)	Men (3)
Migrant Sending	-0.790	-3.129	-1.912
	(0.289)	(1.492)	(0.828)
	[0.220]	[1.257]	[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 Titiunik, 2014).

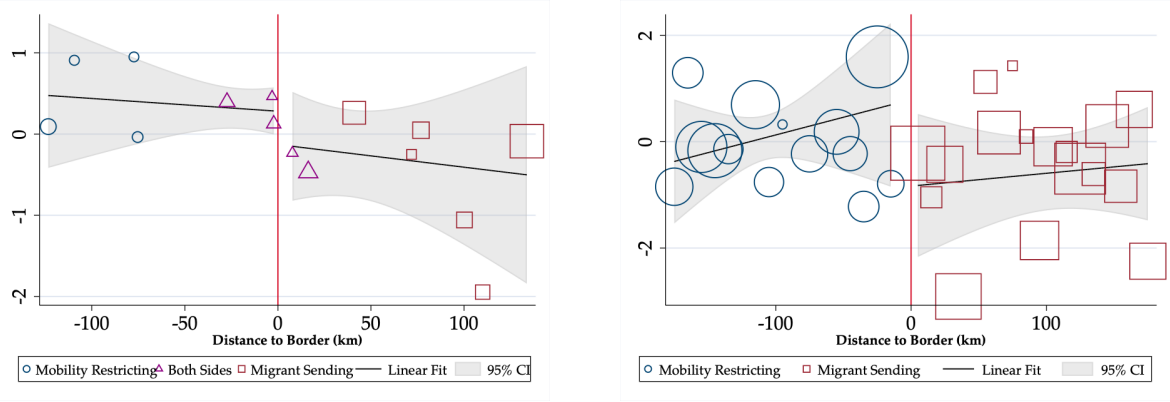
age disparity between a woman and her linked spouse in the census is 0.8 years smaller just inside the former migrant-sending institution. This estimate is meaningful relative to the average of 8.4 years in the former restricted mobility region, and its precision does not seem to be overstated. Figure 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

Figure 9: RD Plots for Age-Disparate Relationships



(a) Census: Women with Spouse

(b) DHS: Men with Last Sexual Partner

Notes: RD plots show the mean age disparity among the specified group in an administrative post (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.

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 7 Column (1) shows that men in the former migrant-sending region are 16 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 7 is consistent with reductions in risk factors associated with age-disparate relationships.

Table 7: HIV Risk Factors Associated with Age Disparities

	Concurrent Partners	*Age at First Sex*	*Condom Used Last Sex*	
	Men	Women	Men	Women
	(1)	(2)	(3)	(4)
Migrant Sending	-0.157	0.813	0.065	0.001
	(0.086)	(0.379)	(0.058)	(0.045)
	[0.081]	[0.403]	[0.061]	[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

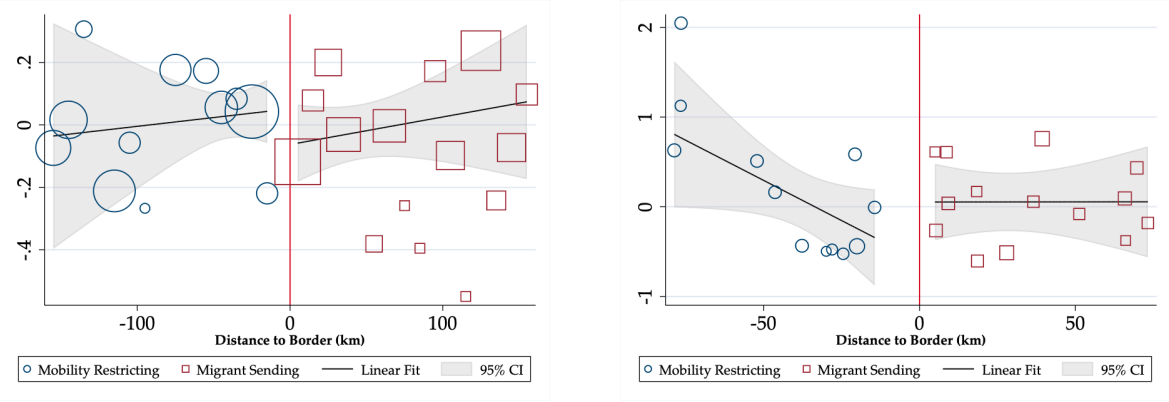
Notes: Standard errors clustered by DHS survey cluster are in parentheses. Regressions estimate a local linear RD specification on each side of the border using a triangular weighting kernel and include age, age squared, longitude, and year fixed effects as controls. Specifications use the MSE-optimal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

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 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 (11) 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

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



(a) Men: Has Concurrent Partners

(b) Women: Age at First Sex

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.

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 my results.

8. Conclusion

Institutions play a major role in economic development (North, 1990) and there is a rich literature on the positive long-run effects of those that promote inclusive prosperity compared to those that simply extract wealth from the population (e.g., Acemoglu, Johnson and Robinson, 2001). But in what is today the developing world, there was never a choice between inclusion and extraction —colonizers instead chose between different types of extractive institutions. In spite of the importance of this decision for understanding the roots of global health and wealth disparities, there is a lack of causal evidence on their comparative impacts.

This paper shows that two extractive institutions common throughout colonial Africa

can have markedly different impacts on HIV prevalence today but result in no long-run differences in economic development. Using the arbitrary border within Mozambique between a migrant-sending and mobility-restricting institution, I find that adults are substantially less likely to be HIV positive just inside the former. The colonial-era and modern evidence suggests that this effect arises from smaller age gaps between spouses and sexual partners in this region, which are consistent with the predictions of a model of a marriage market with bride price after high wages for young men become available.

These results speak to the importance of institutions in shaping present-day outcomes in the developing world. They also show that marriage markets are a channel through historical events affect the present. Additionally, these findings 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 recent era of accelerated globalization. Such lessons are necessary 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 their historical and social roots.

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. [6]

Acemoglu, Daron, and James A. Robinson. 2012. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty.* New York:Crown Business. [2]

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. [2, 6, 42]

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. [11, 12]

Alsan, Marcella. 2015. "The Effect of the TseTse Fly on African Development." *American Economic Review*, 105(1): 382–410. [52]

Alsan, Marcella, and Marianne Wanamaker. 2018. "Tuskegee and the Health of Black Men." *Quarterly Journal of Economics*, 133(1): 407–455. [7]

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. [7, 41]

Ansell, Nicola. 2001. "'Because It's Our Culture!' (Re)negotiating the Meaning of Lobola in Southern African Secondary Schools." *Journal of Southern African Studies*, 27(4): 697–716. [25]

Ashraf, Nava, Natalie Bau, Nathan Nunn, and Alessandra Voena. 2020. "Bride Price and Female Education." *Journal of Political Economy*, 128(2): 591–641. [7]

Ashraf, Quamrul, and Oded Galor. 2011. "Dynamics and Stagnation in the Malthusian Epoch." *American Economic Review*, 101(5): 2003–2041. [24, 26, 27]

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. [13, 31]

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. [6]

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. [15]

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. [7]

Black, Dan A., Seth G. Sanders, Evan J. Taylor, and Lowell J. Taylor. 2015. "The Impact of the Great Migration on Mortality of African Americans: Evidence from the Deep South." *American Economic Review*, 105(2): 477–503. [7]

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. [7]

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica*, 82(6): 2295–2326. [16, 19, 34, 35, 37, 39, 41, 55, 56]

Calvi, Rossella. 2020. "Why Are Older Women Missing in India? The Age Profile of Bargaining Power and Poverty." *Journal of Political Economy*, 128(7): 2453–2501. [7]

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. [34]

Cattaneo, Matias D., Nicolás Idrobo, and Rocío Titiunik. 2019. *A Practical Introduction to Regression Discontinuity Designs: Foundations.* Cambridge, UK:Cambridge University Press. [17]

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. [41]

Chiappori, Pierre-André, Murat Iyigun, and Yoram Weiss. 2009. "Investment in Schooling and the Marriage Market." *American Economic Review*, 99(5): 1689–1713. [7]

Clarence-Smith, Gervase. 1985. *The Third Portuguese Empire, 1825-1975: A Study in Economic Imperialism.* Manchester:Manchester University Press. [9, 10]

Colella, Fabrizio, Rafael Lalive, Seyhun Orcan Sakalli, and Mathias Thoenig. 2020. "Inference with Arbitrary Clustering." [18]

Conley, T.G. 1999. "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics*, 92(1): 1–45. [18]

Corno, Lucia, Nicole Hildebrandt, and Alessandra Voena. 2020. "Age of Marriage, Weather Shocks, and the Direction of Marriage Payments." *Econometrica*, 88(3): 879–915. [5, 7, 24, 28]

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. [53]

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. [52]

Darch, Colin. 1983. "Notas sobre Fontes Estatísticas Oficiais Referente à Economia Colonial Moçambicana: Uma Crítica Geral." *Estudos Moçambicanos*, 4: 103–125. [14]

Dell, Melissa. 2010. "The Persistent Effects of Peru's Mining Mita." *Econometrica*, 78(6): 1863–1903. [6]

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. [6]

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. [5, 31, 35]

Derenoncourt, Ellora. 2021. "Can You Move to Opportunity? Evidence from the Great Migration." [7]

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. [4, 7]

Direcção Provincial dos Serviços de Estatística. 1966. *III Recenseamento Geral da População na Província de Moçambique.* Lourenço Marques:Instituto Nacional de Estatística. [14]

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. [41]

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. [7]

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. [6, 40]

Fernández Cebrián, Pablo. 2021. "Withdrawal of the State: The Provision of Primary Schooling in Mozambique under the Indigenato." [4, 13]

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. [9]

Fuller, Charles Edward. 1955. "An Ethnohistoric Study of Continuity and Change in Gwambe Culture." PhD diss. Northwestern University, Evanston, IL. [13]

Galor, Oded, and David N. Weil. 2000. "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond." *American Economic Review*, 90(4): 806–828. [24, 27]

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. [17]

Gengenbach, Heidi. 2010. *Binding Memories: Women as Makers and Tellers of History in Magude, Mozambique.* New York:Columbia University Press. [14]

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. [9]

Greenwood, Jeremy, Nezih Guner, and Guillaume Vandenbroucke. 2017. "Family Economics Writ Large." *Journal of Economic Literature*, 55(4): 1346–1434. [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. [18]

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. [13]

Guthrie, Zachary Kagan. 2018. *Bound for Work: Labor, Mobility, and Colonial Rule in Central Mozambique, 1940-1965.* Charlottesville, VA:University of Virginia Press. [11, 12]

Harries, Patrick. 1983. "Labour Migration from Mozambique to South Africa; with Special Reference to the Delagoa Bay Hinterland, c. 1862 to 1897." PhD diss. School of Oriental and African Studies, University of London, London. [3, 13]

Harries, Patrick. 1994. *Work, Culture, and Identity: Migrant Laborers in Mozambique and South Africa, c. 1860-1910.* Portsmouth, NH:Heinemann. [3, 10, 12]

Harris, Marvin. 1959. "Labour Emigration among the Moçambique Thonga: Cultural and Political Factors." *Africa*, 29(1): 50–66. [9]

Harrison, Graham. 1998. "Quantos Somos? The Second National Population Census of Mozambique." *Review of African Political Economy*, 25(75): 124–132. [14]

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. [14]

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. [12]

Iliffe, John. 2006. *The African AIDS epidemic: A History.* Athens, OH:Ohio University Press. [7, 14, 31]

Instituto Nacional de Estatística, and ICF. 2019. *Moçambique Inquérito Nacional sobre Indicadores de Malária.* Rockville, MD. [33]

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. [13]

Junod, Henri A. 1912. *The Life of a South African Tribe.* Vol. 1, London:David Nutt. [3, 9, 12]

Kelly, Morgan. 2020. "Understanding Persistence." [16, 17]

Khanna, Gaurav, Caroline Theoharides, and Dean Yang. 2020. "Abundance from Abroad: Migrant Income and Long-Run Economic Development." [7, 31]

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. [52]

Lowes, Sara, and Eduardo Montero. 2021*a*. "Concessions, Violence, and Indirect Rule: Evidence from the Congo Free State." [6]

Lowes, Sara, and Eduardo Montero. 2021*b*. "The Legacy of Colonial Medicine in Central Africa." *American Economic Review*, 111(4): 1284–1314. [7, 35]

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. [6, 40]

Maffioli, Elisa M. 2017. "Is Traditional Male Circumcision Effective as an HIV Prevention Strategy? Evidence from Lesotho." *PLOS ONE*, 12(5): e0177076. [41]

Michalopoulos, Stelios, and Elias Papaioannou. 2014. "National Institutions and Subnational Development." *Quarterly Journal of Economics*, 129(1): 151–213. [6]

Michalopoulos, Stelios, and Elias Papaioannou. 2016. "The Long-Run Effects of the Scramble for Africa." *American Economic Review*, 106(7): 1802–1848. [6]

Ministério de Saúde, Instituto Nacional de Estatística, and ICF. 2018. *Moçambique Inquérito de Indicadores de Imunização, Malária e HIV/SIDA 2015.* Maputo. [33]

Ministério de Saúde, Instituto Nacional de Estatística, and ICF International. 2013. *Moçambique Inquérito Demográfico e de Saúde 2011.* Calverton, MD. [33]

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. [33]

Ministério do Ultramar. 1959. "Carta de Moçambique." [15]

Minnesota Population Center. 2020. *Integrated Public Use Microdata Series, International: Version 7.3.* Minneapolis, MN:IPUMS. [53]

Moran, P. A. P. 1950. "Notes on Continuous Stochastic Phenomena." *Biometrika*, 37(1-2): 17–23. [18]

Morier-Genoud, Éric. 2019. *Catholicism and the Making of Politics in Central Mozambique, 1940-1986.* Rochester:University of Rochester Press. [4, 12]

Mozambique News Agency. 1999. "Huge Growth at Vilankulo Airport." *AIM Reports*. [57]

Mozambique News Agency. 2000. "Tourism Brings in $32 Million." *AIM Reports*. [57]

Mozambique News Agency. 2011. "President Inaugurates New Terminal at Vilankulo Airport." *AIM Reports*. [57]

Murdock, George Peter. 1959. *Africa: Its Peoples and Their Culture History.* New York:McGraw-Hill Book Company. [18, 19]

Méndez-Chacón, Esteban, and Diana Van Patten. 2021. "Multinationals, Monopsony, and Local Development: Evidence from the United Fruit Company." [6]

Newitt, Malyn. 1995. *A History of Mozambique.* Bloomington, IN:University of Indiana Press. [8, 9, 10, 12, 13]

North, Douglass. 1990. *Institutions, Institutional Change and Economic Performance.* New York:Cambridge University Press. [42]

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. [10]

Quinn, Thomas C., Maria J. Wawer, Nelson K. Sewankambo, David Serwadda, Chuanjun Li, Fred Wabwire-Mangen, Mary O. Meehan, Thomas Lutalo, and Ronald H. Gray for the Rakai Project Study Group. 2000. "Viral Load and Heterosexual Transmission of Human Immunodeficiency Virus Type 1." *New England Journal of Medicine*, 342: 921–929. [31]

Reniers, Georges, and Rania Tfaily. 2008. "Polygyny and HIV in Malawi." *Demographic Research*, 19(53): 1811–1830. [41]

Repartição Nacional de Estatística. 1942. *Censo da População em 1940.* Vol. 2-5, Lourenço Marques:Imprensa Nacional de Moçambique. [14]

Reynoso, Ana. 2021. "Polygamy, Co-Wives' Complementarities, and Intrahousehold Inequality." [7]

Rita-Ferreira, A. 1960. "Labour Emigration among the Moçambique Thonga: Comments on a Study by Marvin Harris." *Africa*, 30(2): 141–152. [9]

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. [12]

Saldanha, F. 1940. "Carta da Colónia de Moçambique: Segundo a Sua Actual Divisão Administrativa." [15]

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. [5, 31, 40]

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. [52]

Siow, Aloysius. 1998. "Differential Fecundity, Markets, and Gender Roles." *Journal of Political Economy*, 106(2): 334–354. [5, 25]

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. [8]

Stichter, Sharon. 1985. *Migrant Laborers.* Cambridge, UK:Cambridge University Press. [7]

Tanser, Frank, Till Bärnighausen, Lauren Hund, Geoffrey P. Garnett, Nuala McGrath, and Marie-Louise Newell. 2011. "Effect of Concurrent Sexual Partnerships on Rate of New HIV Infections in a High-Prevalence, Rural South African Population: A Cohort Study." *The Lancet*, 378(9787): 247–255. [31]

Tertilt, Michèle. 2005. "Polygyny, Fertility, and Savings." *Journal of Political Economy*, 113(6): 1341–1371. [5, 7, 24, 25, 26]

Transvaal Chamber of Mines. 1946. "Map of Southern Africa Showing the Recruiting Systems of the Native Recruiting Corporation and the Witwatersrand Native Labour Association." [10]

Vail, Leroy. 1976. "Mozambique's Chartered Companies: The Rule of the Feeble." *Journal of African History*, 17(3): 389–416. [8]

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. [35]

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. [5, 31]

Weinstein, Jeremy M. 2006. *Inside Rebellion: The Politics of Insurgent Violence.* New York:Cambridge University Press. [13]

Wilson, Francis. 1972. *Labour in the South African Gold Mines, 1911-1969.* New York:Cambridge University Press. [10]

World Bank. 2020. *World Slope Model in Degrees.* The World Bank Group. [52]

Yang, Dean. 2008. "International Migration, Remittances and Household Investment: Evidence from Philippine Migrants' Exchange Rate Shocks." *Economic Journal*, 118(528): 591–630. [4]

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. [52]

Appendix A. Data Sources and Variable Definitions

Geographic Traits [19]

- *Elevation*: Average altitude in meters in the 0.25 × 0.25 degree cell. Data from Danielson and Gesch (2011).
- *Rainfall*: Average precipitation in millimeters in the 0.25 × 0.25 degree cell from 1891 to 2016. Data from Schneider et al. (2020).
- *Slope*: Average slope in degrees in the 0.25 × 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 × 0.25 degree cell. Data from Zabel, Putzenlechner and Mauser (2014).

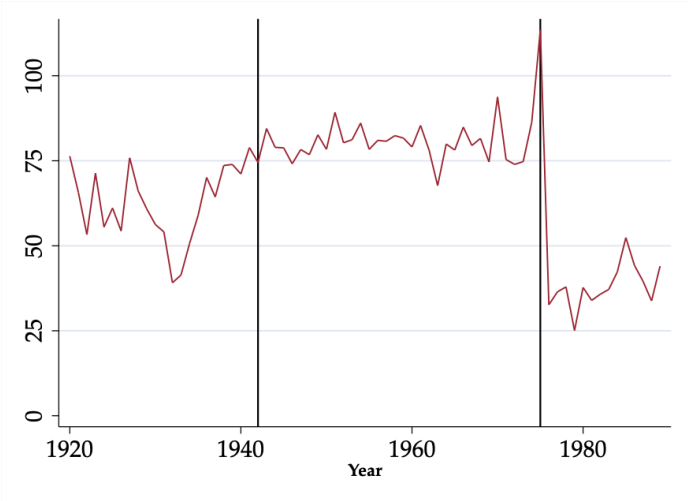
Disease Environment [19]

- *Malaria*: Average malaria transmission stability index value in the 0.25 × 0.25 degree cell. Data from Kiszewski et al. (2004).
- *TseTse*: Average tsetse fly suitability index value in the 0.25 × 0.25 degree cell. Data from Alsan (2015).

Appendix B. Additional Figures

B1. Annual Numbers of Witwatersrand Mine Workers from Southern Mozambique

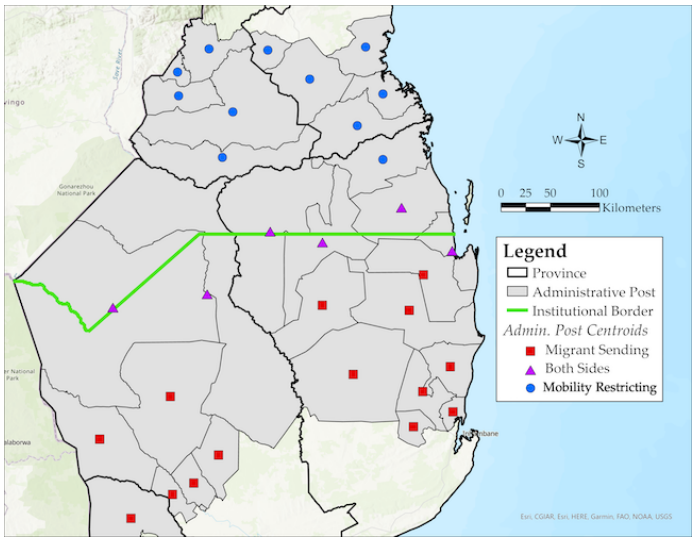
Figure B1: Southern Mozambican Men Received by Mines (000s), 1920-89 [11, 13]



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

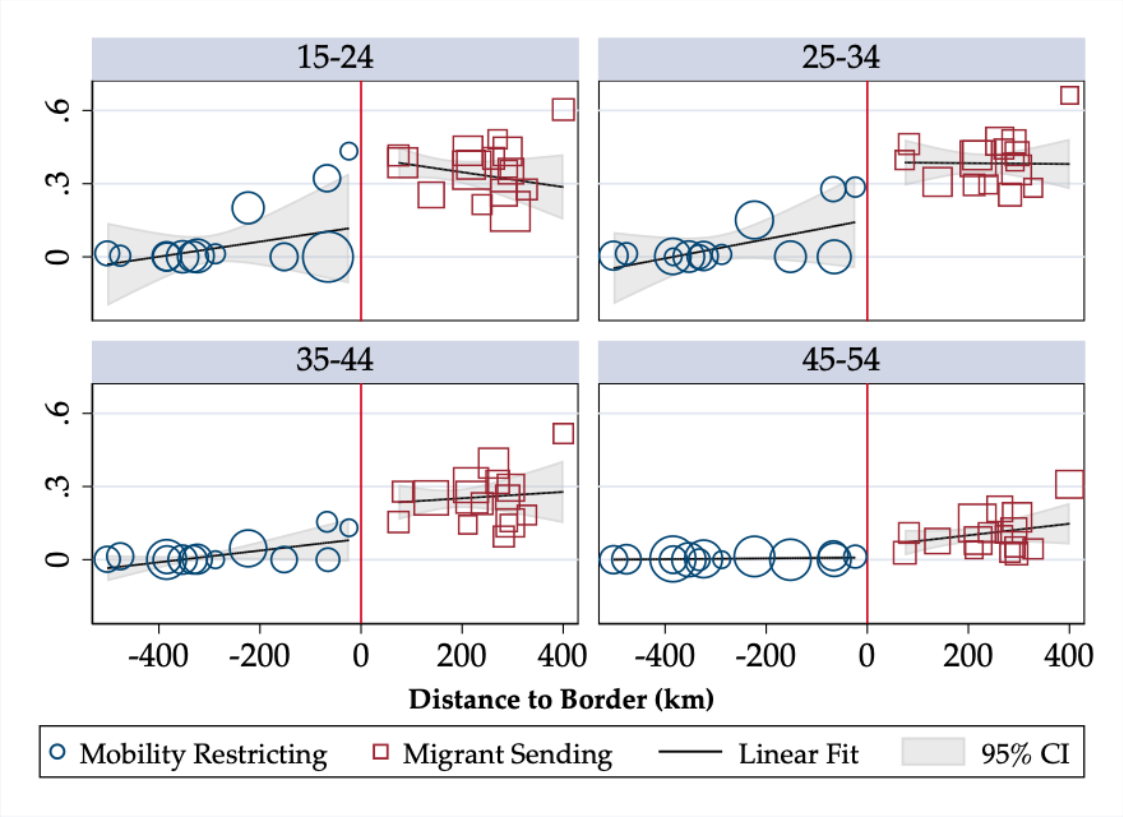
Figure B2: Map of Administrative Posts [18, 38]



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 [20]



Appendix C. Additional Tables

C1. HIV Blood Test Refusals

Table C1: HIV Blood Test Refusals [34]

	Refused Blood Test		
	Pooled (1)	Women (2)	Men (3)
Migrant Sending	0.009	0.010	-0.006
	(0.006)	(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 [41]

	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)
Migrant Sending	0.061	0.027	-0.074	0.017	-0.024	0.015	-0.057	0.031
	(0.028)	(0.014)	(0.065)	(0.049)	(0.043)	(0.081)	(0.094)	(0.175)
	[0.030]	[0.017]	[0.076]	[0.057]	[0.030]	[0.062]	[0.121]	[0.079]
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 p	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 6.1, 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 200,00 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.