History and Human Capital in Developing Economies

by

Jonathan Daniel Denton-Schneider

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in The University of Michigan
2022

Doctoral Committee:

Professor Hoyt Bleakley, Chair
Associate Professor Achyuta Adhvaryu
Assistant Professor Eduardo Montero, University of Chicago
Professor Dean Yang

Jon Denton-Schneider
jdds@umich.edu
ORCID ID 0000-0002-9868-3238

ACKNOWLEDGEMENTS

Completing this dissertation would not have been possible without the thoughtful guidance, invaluable feedback, and endless patience of my committee members Hoyt Bleakley, Eduardo Montero, Ach Adhvaryu, and Dean Yang. I am hugely indebted to them for being exceptional mentors and always pushing me to become a better economist. My goal as an academic is to emulate the inventiveness and rigor of their research as well as their graciousness as advisors and colleagues. Learning from them was a privilege.

I owe special thanks to Hoyt and Eduardo for their incredibly generous investments of time in me that I hope yields the returns they deserve. My conversations with Hoyt began as soon as I arrived in Ann Arbor; he taught me the importance of seeking out critical feedback early and often, and how to deliver it with wit and humor. Martha Bailey once told me that I sounded like Hoyt when I described my approach to a project, which was the highest professional compliment I have received. Eduardo's support and encouragement were the key to building my self-confidence as a researcher, saving me numerous times from letting setbacks dissuade me from continuing with a project—and even the doctoral program itself. I deeply value his friendship, collaboration, and wisdom, although the latter is called into question by his being a Liverpool fan.

Along with my committee members, there were several other professors and staff at Michigan who made it possible for me to arrive at this point. Ach and Tanya Rosenblat were extraordinary third-year paper readers: they dedicated great effort and many hours to steering me through the process of conducting research, even though I was taking a risk on a field and region that I had no experience with. Sara Heller and David Miller provided the world-class instruction and welcoming learning environment that helped me overcome my previous struggles in and debilitating self-doubt regarding econometrics and microeconomic theory. Dropping by Laura Flak's office to chat about pets and steal her candy kept me grounded in my third year, and Miriam Rahl's support and kindness helped me navigate my fourth through sixth years.

During that period, my research was supported by a National Institute on Aging training grant to the Michigan Population Studies Center (T32AG000221). This traineeship

changed my career with generous research support, including a grant from the Marshall Weinberg Endowment Fund, as well as protected time to work on my dissertation. Throughout my time at Michigan, my aunts Laura and Susan Jacobson let me live in their aunt Marjorie Jacobson's home, which was unbelievably generous and leaves me in their eternal debt. (But I did look after the house, so maybe that reduces it a little bit!)

In addition, there were professors, coaches, and teachers from before my arrival at Michigan whose dedication and advocacy were instrumental in what I have accomplished since they changed my life, and I am forever grateful to them for their care during the three most difficult periods of my life. At Arizona, Price Fishback and Tauhid Rahman sparked in me passions for economic history and development economics, and the steadfastness of their belief in my decision to pursue a doctorate helped me persevere through the rejection of all but one of my applications and the lows of the first year. My coaches Brandy Maben and Don Swartz taught me that there was so much more to who I am than how fast I (never quite managed to) swim, to never to take myself too seriously, and that our greatest power in life is the power to choose. And at Marin Academy, Derek Anderson showed me how to see broad historical trends and entire cultures through the lens of something smaller, as well as the value of sticking to one's principles and objective standards when the prevailing winds blow in the opposite direction.

I also could not have done without the love and support of the lifelong friends I made in Michigan, Arizona, and LGBTQ sports. James and Joye Allen, Jennifer Mayo and Luke Burns, Michael and Emily Bertsche Murto, Ben Pyle and Gillian Wener, Jawad Bandar, Scott Baty, Wade Rakes and Nick Miller, Melissa Magdefrau, Laura Taylor, and Matt Hatzke and Holden Davis inspire me to be more caring, more open to adventure, and even more hilarious than I already am. I particularly thank Jennifer and Laura for helping me survive the job market, James and Melissa for making our time in classes together fun, and the four of them for strangely wonderful study sessions. Jeff and Gaby Sandoval Amlee, Zac Hojnacki, and Austin Ringquist have made me a better person and have given me more second chances than I deserve, which I can never repay them for. I am so excited to see everyone's careers taking off, their children growing up, and their pets getting "fluffier." Their friendship is a blessing and that I somehow earned it is a miracle.

Finally, I come to the impossible task of expressing in words my gratitude to my family. To put it simply, I would be nowhere without them, and everything I am is because of them. My parents Jane Denton and Jeff Schneider, my brothers Ben and Sam Denton-Schneider and sister-in-law Kate Varandas, and my fiancé Zach Yetmar have shown me what it means to love and be loved, and to aim for both excellence and goodness. I hope to always make them proud and to be worthy of their sacrifices.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

Appendix

ABSTRACT

This dissertation studies the historical causes and economic consequences of poor health—an important component of human capital—in developing countries, as well as the policies that can remedy its negative impacts. The link between these chapters is infectious disease: the first examines it as an outcome of historical processes, the second investigates links between childhood morbidity and the likelihood of contracting a deadly illness as a young adult, and the third studies novel domains in which disease control can have short- and long-run impacts on a developing economy.

Chapter 1 shows how common colonial labor regimes in Africa shaped HIV prevalence in Mozambique today through their long-run impacts on marriage and dating practices. To compare two of these extractive institutions—one pushing men into circular migration and one restricting their mobility—I exploit the arbitrary border within Mozambique that separated the regions under them for a half-century (1893-1942). Consistent with historians' accounts, spousal age gaps in the colonial era were smaller in the migrant-sending area, even after the border was erased and circular migration rates converged. Today, HIV prevalence is substantially lower in this region, likely due to the narrower age gaps between partners that reduce exposure to the virus, but development outcomes are similar. These results demonstrate 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.

Chapter 2 shows that treating schistosomiasis in secondary school-age girls in Zimbabwe reduced their likelihood of contracting HIV as young women by changing their marriage and dating behaviors. I examine Zimbabwe's nationwide school-based deworming program (2012-17) that substantially reduced rates of the urogenital form of the disease. Using a difference-in-differences design, I find that 3 years after deworming began, HIV prevalence among young women had fallen more in formerly high-schistosomiasis districts than in low-morbidity ones. Human capital's effects on marriage markets appear to explain the results: these young women were more likely to still be in school, age gaps with their partners shrank, and HIV risk factors associated with age gaps in relationships decreased more as well. These results imply that a cheap

treatment for a common childhood disease can also be a highly cost-effective method of combating one of the modern world's deadliest pandemics.

Chapter 3 shows that Brazil's efforts to eliminate Chagas Disease transmission had short- and long-run effects in domains that are important for economic development but have not yet been included in cost-benefit analyses of disease control. Using a difference-in-differences strategy, we make comparisons before and after the campaign (1984-89) across states and municipalities with varying levels of pre-treatment vector prevalence. We find that adults' employment rates rose shortly after spraying began and cohorts treated as children had higher incomes as adults, and the latter effect was much larger for non-white Brazilians. Using a triple-differences strategy, we also show that Brazil's government-run health care system—which consumes 4 percent of GDP—spent substantially less on hospitalizations due to circulatory diseases than other causes. These results imply that the benefits of disease control can be larger than previously assumed, and that combating neglected tropical diseases can help to speed convergence in societies with large racial disparities.

CHAPTER I

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

1.1 Introduction

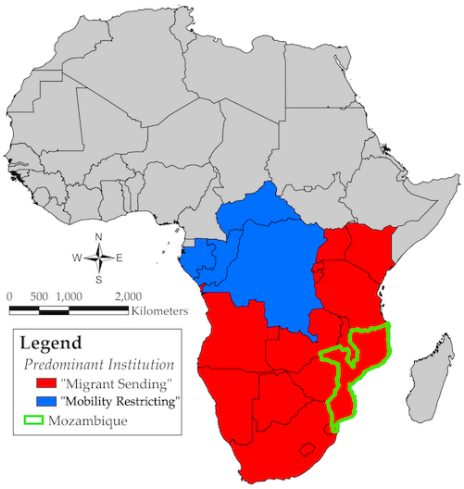
After European powers partitioned Africa at the end of the nineteenth century, they began extracting wealth from their colonies by imposing several types of institutions. Amin (1972) underscored the significance of these regimes in African history and development by grouping present-day countries into "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 1.1 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

[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 1.1: 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.

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 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 et al., 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 1.2, and Figure 1.2 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 1.3, 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.

In Section 1.4, 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.

[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 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 1.5 an overlapping generations model 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 et al., 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

[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).

[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.

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.

As I show in Section 1.6, 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 1.7 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 et al., 2001; Banerjee and Iyer, 2005; Dell, 2010; Michalopoulos and Papaioannou, 2014, 2016), and there is recent evidence specifically on concessions (Dell and Olken, 2020; Lowes and Montero, 2021a; Méndez-Chacón and Van Patten, 2021). However, we know little about the short- or long-run impacts of the colonialist's choice from a menu of extractive institutions. My contribution in this area is to provide the first evidence on this question,

[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.

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 et al., 2019; Derenoncourt, 2021; Khanna et al., 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.

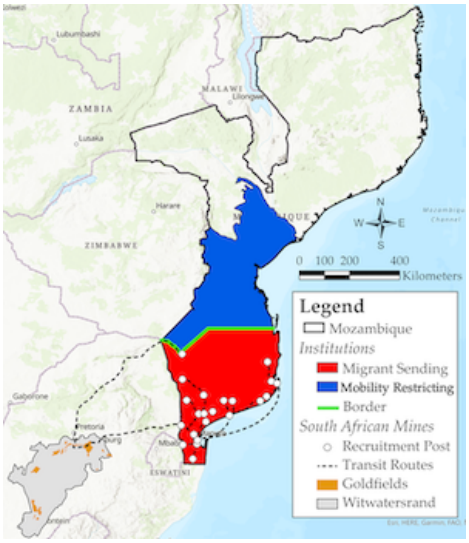
In addition, the economic analysis of non-Western marriage markets—which determine how most of the world marries—is an expanding area of study (Tertilt, 2005; Corno et al., 2020; Reynoso, 2021). I contribute in this field by showing how they interact with labor market conditions to shape behaviors affecting human capital in the long run (Chiappori et al., 2009; Greenwood et al., 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, 2021b). 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.

1.2 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

Figure 1.2: 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).

to the end of its civil war and the explosion of HIV in the late 1990s.

1.2.1 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 1.2 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:

[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).

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

1.2.2 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.

1.2.2.1 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 durations were limited to one year with a possible six-month extension and a mandated

[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).

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 infrastructure to move them from there to the Witwatersrand. Figure 1.2 shows this network in 1946 and Appendix A.2.1 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.

1.2.2.2 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

[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).

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).

1.2.3 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.

Figure 1.3: 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.

(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).

1.2.4 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 1.3 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

[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).

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 A.2.1) 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]

1.3 Colonial Data and Empirical Strategy

1.3.1 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).

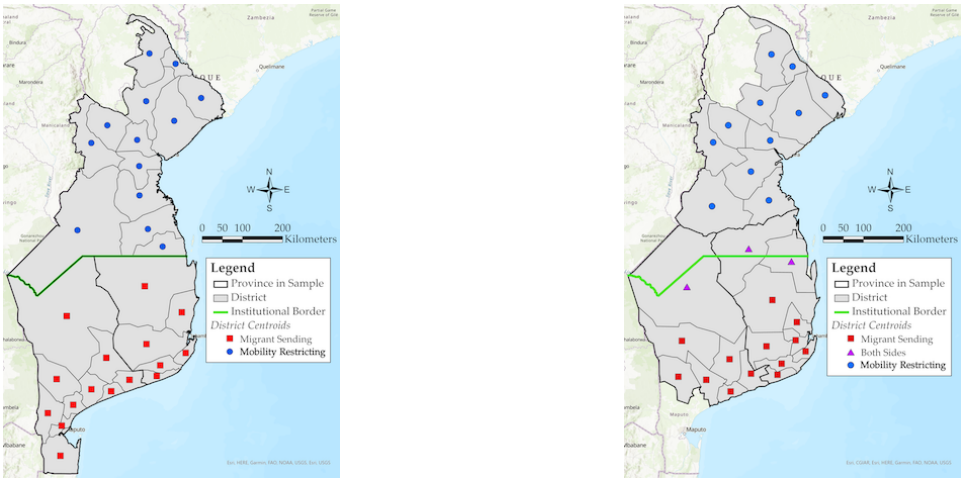
Similarly, the 1960 data allow for the most reliable and longest-run comparison of the two regions during the colonial period. This census took place 18 years after the mobility-restricting institution ended and it was the last one before the Mozambican War of Independence (1964-74). As such, it does not suffer from the data collection problems that can arise when governments participate in internal conflicts (Barakat et al., 2002).

1.3.1.1 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 1.2, 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.

[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 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).

Figure 1.4: 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.

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]

1.3.1.2 Georeferenced Sample

I match these district-level data to administrative maps of Mozambique from each year (Saldanha, 1940; Ministério do Ultramar, 1959). Figure 1.4 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]

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

[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.

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

1.3.2 Empirical Strategy

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

$$y_d = \alpha + \tau \text{MigrantSending}_d + f(\text{Distance}_d) + \text{Lon}_d + \epsilon_d \quad \text{for } d \in B, \tag{1.3.1}$$

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 capture east-west trends.[23] I use a local linear specification estimated separately on each side with a triangular kernel (Cattaneo et al., 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 (1.3.1) 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.

1.3.2.1 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

[22] Because the colonial data are reported at the district level, there are too few centroids near the border to estimate the Calonico et al. (2014) mean squared error (MSE) optimal bandwidth.

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

errors may be too small due to similarity among neighbors. Intuitively, assuming the statistical independence of observations would overstate the information each one adds to the estimation, leading to inflated precision.

I take two steps to address this potential problem. First, I calculate Conley standard errors allowing for arbitrary spatial correlation between observations within 100 km of each other, imposing a linear decay in relationships over this bandwidth ("Bartlett kernel") (Conley, 1999; Colella et al., 2020). I report these standard errors in addition to those robust to heteroskedasticity. Second, I diagnose positive spatial autocorrelation in the residuals by computing the Moran (1950) I-statistic, which is the slope of the line in a weighted regression of neighbors' values on each unit's value.[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]

1.3.3 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 1.5 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 (1.3.1) clustering standard errors by third-level administrative unit ("administrative posts," shown in Appendix A.2.2).[27] Consistent with the border being arbitrary, Table 1.1 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 1.6.2 for details on RD bandwidth selection when using geographically disaggregated data.

Figure 1.5: Ethnic Group Homelands in Mozambique

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

Table 1.1: 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	182.693	40.936	0.179	49.189	10.838	1.259
Mobility Restricting SD	109.318	44.759	0.132	9.894	1.757	0.078

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 et al., 2014). Data sources and variable definitions are in Appendix A.1.

1.4 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 1.2 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 1.6 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.

1.4.1 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 1.6a 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 A.2.3 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. This pattern implies that the regime significantly constrained men's mobility and its end in 1942 led to major changes.

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.

1.4.2 Marriage and Fertility

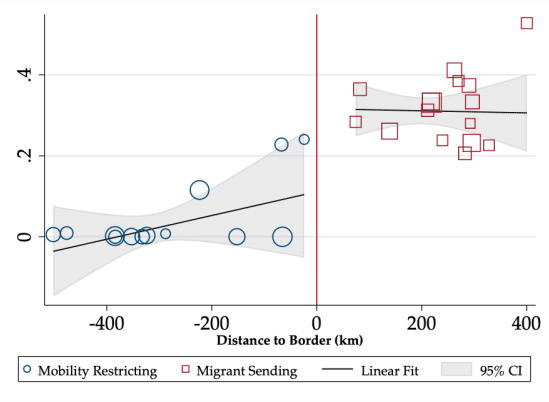
I then turn to comparing marriage market outcomes across the institutional border, as narrative histories argue that earning bridewealth was an important motivation for men

Table 1.2: Comparing Institutions during the Colonial Era

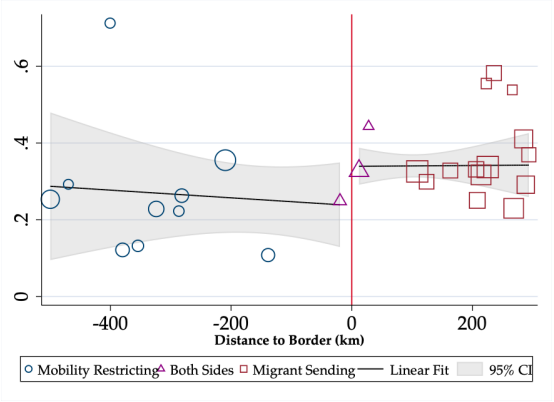
	Panel A: 2 Years before End of Mobility Restricting Institution (1940)					
	Migration	*Marriage and Fertility Ratios*			*Human Capital*	
	Men Work Abroad (1)	Marriage Ages 15-24 (2)	Marriage Ages 25-34 (3)	Child-Woman (4)	Boys in School (5)	Girls in School (6)
Migrant Sending	0.207	0.269	0.312	0.202	-0.028	-0.001
	(0.089)	(0.080)	(0.113)	(0.098)	(0.016)	(0.006)
	[0.085]	[0.061]	[0.088]	[0.092]	[0.015]	[0.006]
Observations	29	28	28	28	29	29
Bandwidth	-503, 401	-503, 401	-503, 401	-503, 401	-503, 401	-503, 401
Spatial Autocorrelation	-0.14	0.26	0.14	-0.15	-0.03	-0.06
Spatial Autocorrelation SD	0.11	0.11	0.11	0.12	0.11	0.10
Mobility Restricting Mean	0.047	0.364	0.696	0.848	0.050	0.006
Mobility Restricting SD	0.089	0.177	0.304	0.144	0.044	0.009
	Panel B: 18 Years after End of Mobility Restricting Institution (1960)					
	Migration	*Marriage and Fertility Ratios*			*Human Capital*	
	Men Work Abroad (1)	Marriage Ages 15-24 (2)	Marriage Ages 25-34 (3)	Child-Woman (4)	Boys in School (5)	Girls in School (6)
Migrant Sending	-0.025	0.102	0.106	0.056	-0.014	-0.026
	(0.049)	(0.057)	(0.147)	(0.086)	(0.033)	(0.024)
	[0.050]	[0.053]	[0.131]	[0.082]	[0.029]	[0.021]
Observations	27	27	27	28	28	28
Bandwidth	-500, 294	-500, 294	-500, 294	-500, 294	-500, 294	-500, 294
Spatial Autocorrelation	-0.09	-0.03	-0.15	0.05	0.44	0.24
Spatial Autocorrelation SD	0.12	0.12	0.10	0.12	0.12	0.13
Mobility Restricting Mean	0.163	0.267	0.635	0.807	0.089	0.041
Mobility Restricting SD	0.116	0.166	0.235	0.123	0.061	0.040

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 1.3. 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.

Figure 1.6: 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 1.3.

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 1.2 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.[28] The point estimates are large relative to ratios in the mobility-restricting institution for ages 15 to 24 (averaging around 1:3) and 25 to 34 (2:3). Panel B Columns (3) and (4) show that these differences decreased but remained sizable after the mobility-restricting regime ended, though the estimate for ages 25 to 34 is imprecise. The RD plot in Figure 1.6b 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 1.5.

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.[29] The positive response of fertility to circular migration (and

[28]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.

[29]The estimate excludes a district just inside the migrant-sending region with a child-woman ratio of 1.3,

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.

1.4.3 Human Capital Accumulation

Finally, I examine whether differences between the institutions affected investments in children's human capital. Table 1.2 Panel A Column (6) shows that prior to the end of the mobility-restricting institution—after Protestant missionaries were expelled from the migrant-sending region but before Catholic missions assumed responsibility for education there—boys just inside the migrant-sending region were 3 p.p. less likely to be in school. This effect is large relative to the 5-percent enrollment rate for boys in the mobility-restricting institution. In contrast, there is no effect for girls in Column (7).

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

1.5 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 et al. (2020) as well as the Malthusian economies in Galor and Weil

far higher than any other district in the sample. Its exclusion makes the estimated coefficient smaller.

(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.[30]

1.5.1 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.[31] After individuals 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.[32] The discount factor is $\beta \in (0, 1)$.

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.[33] 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.[34] At the end of the period, men pay bride prices p_t to their in-laws.[35]

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

[30] 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.

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

[32] This asymmetry makes fecund women scarce (Siow, 1998)

[33] This term could reflect the greater discomfort from arduous physical labor in old age.

[34] 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.

[35] In Southern Africa, bride prices can be paid after a marriage occurs (Ansell, 2001).

historical context discussed in Section 1.2.

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 $\frac{\psi(n_t^A)^2}{w_t^A}$.[36] After they are born, households consume, all old adults and the fixed share of young adults die, and the period ends.

1.5.1.2 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.[37]

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.[38] 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

$$\max_{w_t^O, n_t^O \geq 0} \log\left[y_t^O + \frac{p_{t-1}}{1-\mu} - w_t^O(p_t - y_t^F) - \frac{\psi(n_t^O)^2}{w_t^O}\right] + \lambda \log(n_t^O), \tag{1.5.1}$$

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.[39] Their problem as young men in t is

[36] Note that this cost increases in the number of children and decreases with the number of wives.

[37] 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.

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

[39] Formally, it requires that this pain outweighs the gains from more consumption and children: $\delta >$

$$\max_{w_t^Y, n_t^Y \geq 0} \log\left(y_t^Y - w_t^Y(p_t - y_t^F) - \frac{\psi(n_t^Y)^2}{w_t^Y}\right) + \beta(1-\mu)\left[\log\left(w_t^Y y_{t+1}^F\right) + \lambda \log(n_t^Y)\right]. \tag{1.5.2}$$

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 (1.5.1) except their share gets scaled up by the fraction who did not marry.

1.5.1.3 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.[40] 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

$$y_t^F = \left(\frac{AX}{N_t + (1-\mu)N_{t-1}}\right)^\alpha \equiv x_t^\alpha. \tag{1.5.3}$$

1.5.1.4 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

$$N_{t+1} = N_t n_t^Y + (1-\mu)N_{t-1} n_t^O. \tag{1.5.4}$$

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

$$N_t = N_t w_t^Y + (1-\mu)N_{t-1} w_t^O. \tag{1.5.5}$$

$\log\left[y_{t+1}^O + w_t^Y y_{t+1}^F - w_{t+1}^O(p_{t+1} - y_{t+1}^F) - \frac{\psi(n_{t+1}^O)^2}{w_{t+1}^O}\right] - \log(w_t^Y y_{t+1}^F) + \lambda[\log(n_t^Y + n_{t+1}^O) - \log(n_t^Y)]$.

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

1.5.2 Predictions

I solve for the equilibria of this marriage market and economy defined by the optimal choices in (1.5.1) and (1.5.2) taking the female wage in (1.5.3) and the bride price p_t as given, population evolving according to (1.5.4), and the marriage market clearing as in (1.5.5). 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.

1.5.2.1 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: $y_0^Y < x_t^\alpha < y_0^O$.[41] The first-order conditions for (1.5.1) yield period-t old men's choices

$$n_t^O = w_t^O \frac{\sqrt{p_t - x_t^\alpha}}{\sqrt{\psi}},\ w_t^O = \frac{\lambda\left(y_0^O + \frac{p_{t-1}}{1-\mu}\right)}{2(1+\lambda)(p_t - x_t^\alpha)}. \tag{1.5.6}$$

Substituting the expressions in (1.5.6) into (1.5.4) and (1.5.5) 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:

$$w_0^O = \frac{1}{1-\mu} = n_0^O \implies p_0 = \left(\frac{AX}{N_0(2-\mu)}\right)^\alpha + \psi. \tag{1.5.7}$$

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 et al. (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.

1.5.2.2 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: $y_0^Y < x_t^\alpha < y_0^O < y_1^Y$.[42] In this new regime, the expressions for period-t old men's

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

[42] 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.

Table 1.3: 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.

choices are the same as in (1.5.6). However, now the ϵ share of period-t young men with high wages make decisions defined by (1.5.2):

$$n_t^Y = \frac{\sqrt{\beta(1-\mu)\lambda w_t^Y}\sqrt{y_1^Y - w_t^Y(p_t - x_t^\alpha)}}{\sqrt{\psi(\beta(1-\mu)\lambda+2)}},\ w_t^Y = \frac{\beta(1-\mu)(\lambda+2)y_1^Y}{2[\beta(1-\mu)(1+\lambda)+1](p_t - x_t^\alpha)}. \tag{1.5.8}$$

Substituting the demands for wives into the market clearing condition (1.5.5) 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 steady-state bride price, the result is a closed-form solution for the period-1 value,

$$p_1 = x_0^\alpha + \epsilon\frac{\beta(1-\mu)(\lambda+2)y_1^Y}{2[\beta(1-\mu)(1+\lambda)+1]} + (1-\mu)\frac{\lambda\left(y_0^O + \frac{x_0^\alpha+\psi}{1-\mu}\right)}{2(1+\lambda)}. \tag{1.5.9}$$

which is clearly greater than p_0 in (1.5.7) 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 1.3, 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.

1.5.3 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.

1.5.3.1 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).

1.5.3.2 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 et al., 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.

1.6 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 1.4 and 1.5 report RD estimates for the respective outcomes, and Figure 1.8 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.

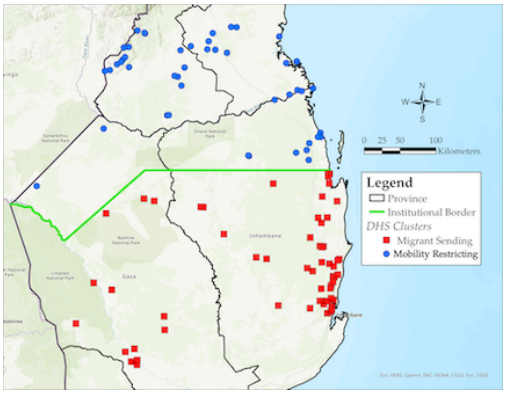
1.6.1 Data

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 1.7 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.[43]

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 A.4, 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.

[43]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.

Figure 1.7: 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 et al. (2010), Ministério de Saúde et al. (2013), Ministério de Saúde et al. (2018), and Instituto Nacional de Estatística and ICF (2019).

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.[44] For economic development, the variables of interest are an index of 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).[45]

1.6.2 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 (1.3.1). I modify it to be:

$$y_{i,c} = \alpha + \tau \text{MigrantSending}_c + f(\text{Distance}_c) + \text{Lon}_c + \mathbf{X}_i\beta + \delta_t + \epsilon_{i,c} \quad \text{for } c \in B^*_{\text{MSE}} \tag{1.6.1}$$

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

[44]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.

[45]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 Calonico et al. (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.

1.6.2.1 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 et al. (2008) recommend. The second concern is spatial autocorrelation, which again I address by calculating Conley standard errors and Moran I-statistics as discussed previously.[46]

1.6.3 Results: HIV Prevalence

I first examine the spatial distribution of HIV among adults along the institutional border. Table 1.4 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 1.8a and 1.8b 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 A.3.1 that there are no substantive differences in blood test refusal rates along the border, helping to rule out selection into testing due to history or other factors (Lowes and Montero, 2021b).

To rationalize the large effect sizes estimated above, I compare the age profiles of HIV prevalence in the two institutions.[47] Specifically, I calculate the mean seroprevalence for each 10-year age group within each sex's MSE-optimal RD bandwidth.[48] Figures 1.8c and 1.8d plot these age profiles. A clear pattern emerges among women: HIV prevalence for

[46] 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.

[47] 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 et al., 2016).

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

Table 1.4: HIV Prevalence

	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
Mobility Restricting SD	0.411	0.411	0.400

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 et al., 2014).

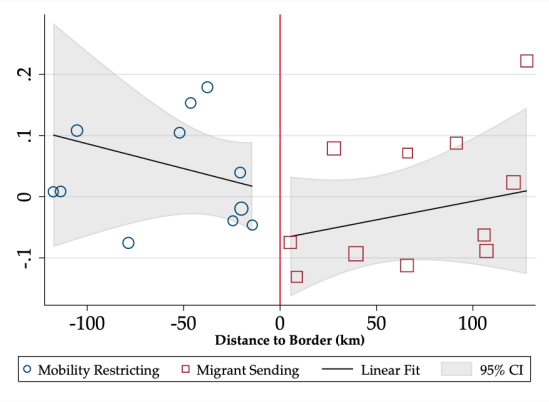
every age group is lower (or at a minimum no greater) in the former migrant-sending institution than in the former restricted mobility region. It is most apparent for women ages 25 to 34, when HIV prevalence peaks in the former mobility-restricting institution. There is a similar pattern for men but 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 1.5. The peak of women's seroprevalence in the former mobility-restricting institution is both larger than the one across the border and farther away from the male peak. This pattern could arise from relationships with wider age disparities, a channel I examine in Section 1.7.

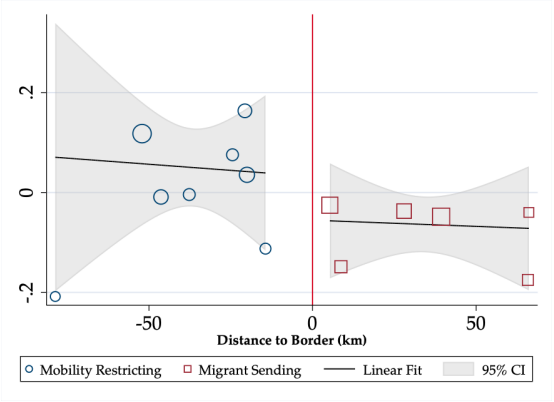
1.6.4 Results: Economic Development

Next, I compare economic development outcomes in the former institutions. The point estimates in Table 1.5 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 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.

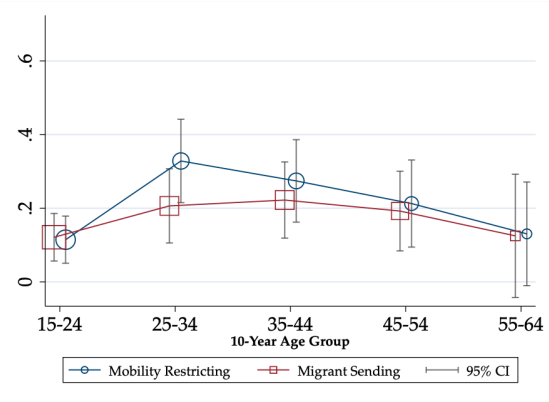
Figure 1.8: HIV Prevalence RD Plots and Age Profiles



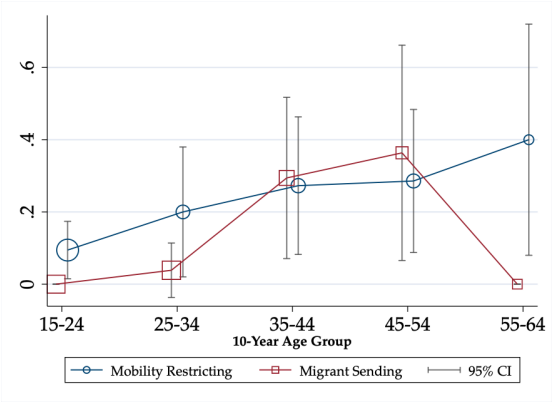
(a) RD Plot: Women



(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.

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.

1.7 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

Table 1.5: 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
Mobility Restricting SD	0.992	0.487	2.565	3.037

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 et al., 2014).

colonial-era marriage market results in Section 1.4 and the age profiles of seroprevalence in Section 1.6. 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.

1.7.1 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.[49]

[49]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

Table 1.6: Age-Disparate Relationships

	Spouse	*Last Sex Partner*	
	Women	Women	Men
	(1)	(2)	(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
Mobility Restricting SD	6.585	5.324	4.167

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 et al., 2014).

However, the census lacks information on sexual activity and the data are at a much coarser geographic resolution (administrative posts, see Appendix A.2.2 for a map).[50]

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.[51] 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.[52]

Table 1.6 reports the results of estimating equation (1.6.1) for age disparities in these datasets and Figure 1.9 presents RD plots for these outcomes. Column (1) shows that the

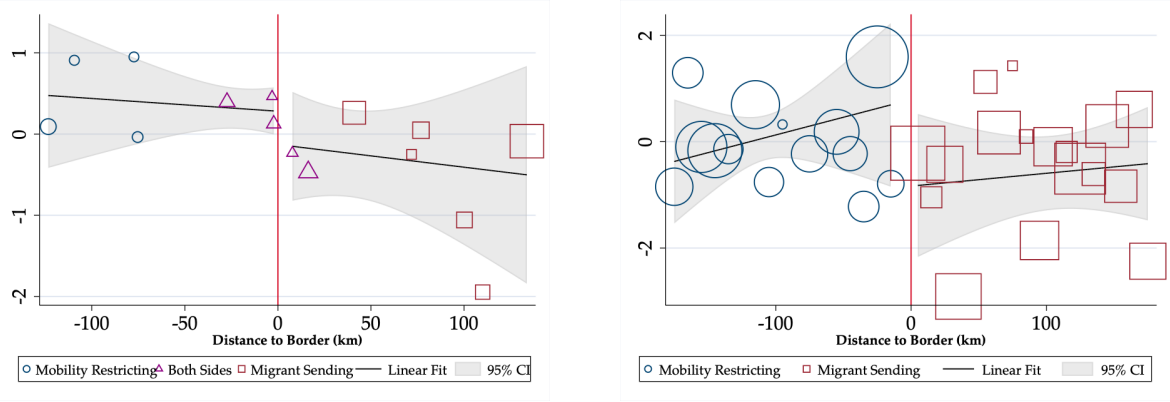
sample is a non-trivial consideration.

[50]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 1.3.2.1, and in the RD plots I take the same approach to the administrative posts with area on both sides of the border.

[51]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.

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

Figure 1.9: 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.

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 1.9a 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 1.9b shows the RD plot for men. Taken together, these results suggest age-disparate relationships are important in explaining the HIV prevalence result.

1.7.2 Risk Factors Associated with Age-Disparate Relationships

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

Table 1.7 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

Table 1.7: 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	16.127	0.069	0.056
Mobility Restricting SD	0.439	2.007	0.256	0.230

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 et al., 2014).

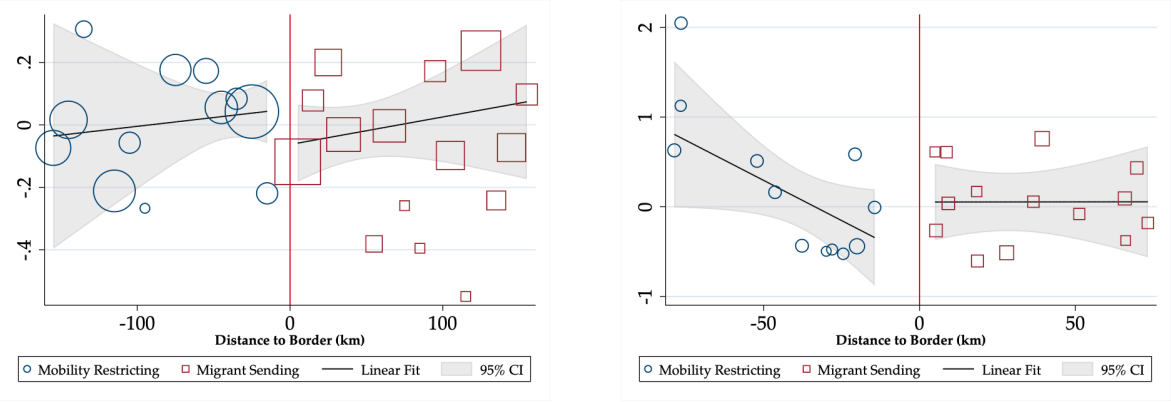
occur 0.8 years later in the former migrant-sending region, which is non-trivial relative the restricted mobility mean of 16.1 years. Although the estimates are somewhat imprecise, the changes at the border are apparent in Figure 1.10, especially for concurrency.

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 1.7 is consistent with reductions in risk factors associated with age-disparate relationships.

1.7.3 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 1.6. 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

Figure 1.10: 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.

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 A.3.2, I estimate equation (1.6.1) for these outcomes. The only ones with effects pointing in the direction of lower HIV prevalence in the former migrant-sending region are fewer women being forced to engage in sex (-2.4 p.p. relative to a restricted mobility mean of 8.7 percent) and more men who are medically circumcised (3.1 p.p. relative to a restricted mobility mean of 19.2 percent), but both estimates are highly imprecise. Additionally, the largest and only precisely estimated effect—6.1 p.p. more women having a genital ulcer in the last year compared to a restricted mobility mean of 0.5 percent—*increases* HIV risk in the former migrant-sending region. As such, these results suggest that the commonly-cited risk factors above do not explain my results.

1.8 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 et al., 2001). But in what is today the developing world, colonizers never chose between inclusion and extraction but rather 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.

CHAPTER II

Childhood Health, Marriage Markets, and Young Women's HIV: Evidence from Deworming in Zimbabwe

2.1 Introduction

Along with its terrible human toll across Sub-Saharan Africa, untreated HIV infections have also imposed devastating economic costs on this region.[1] The advent of antiretroviral therapy (ART), which treats HIV, and its widespread distribution have thus been nothing short of a miracle. But while its costs are far less steep than those of untreated HIV, this miracle still does not come cheaply: in 2019, ART accounted for nearly two-thirds of the more than $3 billion spent on HIV-related commodities in 34 highly-affected countries by their governments and international organizations (Oum et al., 2021). Because people living with HIV must take ART for their entire lives, these expenses will likely remain substantial for decades, even as donor funding to combat the HIV pandemic has plateaued and fallen since the late 2000s (Kates et al., 2020).

Therefore, averting new HIV infections—or even simply *delaying* them for several years—would yield significant savings for governments of high-HIV prevalence countries and international donors.[2] It would be especially true for Eastern and Southern African countries, where more than two-fifths of the world's 1.7 million new HIV infections occurred in 2019 (UNAIDS, 2020). In the case of Zimbabwe, delivering ART costs around

[1] Individuals suffering from AIDS were more often absent from their jobs or out of work entirely (Habyarimana et al., 2010; Bor et al., 2012), pushing young sons into the labor force in their place (Thirumurthy et al., 2008). Also, savings and children's schooling decreased as life expectancy declined, even in HIV-negative households (Baranov and Kohler, 2018). As these impacts reverberated through broader economies, rates of income growth fell across the region (Tompsett, 2020).

[2] With a 5-percent discount rate, the present value of lifelong ART for a 20-year old who lives to age 70 and contracts the virus 5 years from now is one-third lower than if she contracts it today. The savings are even greater if the drugs become cheaper over time, as has been the case for ART. A lag between infection and treatment initiation also raises this percentage, though the amount saved is smaller after discounting.

$175 per patient per year (Benade et al., 2021), or 15 percent of GDP per capita; with 1.23 million HIV-positive adults (almost 13 percent) and 93 percent of them receiving treatment (UNAIDS, 2022), ART costs 1.2 percent of its GDP every year.

By far, the group comprising the largest share of new infections in Zimbabwe and its neighbors—nearly one-third—is women aged 15 to 24 (UNAIDS, 2020).[3] Their risk is so disproportionately high due in large part to common marital practices like age-disparate relationships, which can provide them with economic and psychological benefits (Leclerc-Madlala, 2008). Nonetheless, as older men are some of the most likely to be HIV positive, the result is a cycle of high transmission from older men to younger women to their male peers as these women age (de Oliveira et al., 2017). Unprotected sex is also more common within such relationships, likely as a result of power disparities between partners associated with these age gaps (Tangeni, 2007).

Because these marriage market outcomes for young women can be shaped by their human capital (Becker, 1991), which is in turn shaped by their childhood health (Bleakley, 2010a), can improving the health of girls lower their chances of contracting HIV as young women? If so, does it work by changing the partners they match with and their sexual behaviors? And if the childhood health improvement is an exceptionally cheap one like parasitic worm (helminth) control—as in Miguel and Kremer (2004)—how much would it change the cost-benefit analysis of deworming?

To answer these questions, I study the effects of a nationwide deworming program in Zimbabwe (2012-17), where 8 percent of women aged 15 to 24 were HIV positive when it began. As I describe in Section 2.2, the predominant helminth in the country at that time was *Schistosoma haematobium*, which causes the urogenital form of schistosomiasis (also known as bilharzia), a neglected tropical disease estimated to affect over 100 million people in Sub-Saharan Africa and 230 million worldwide. Infection with *S. haematobium* can result in genital ulcers along with morbidity that limits learning and schooling, and it is most common among adolescents (Colley et al., 2014).[4] Therefore, urogenital schistosomiasis in these ages could affect the HIV status of young women through 3 channels: ulcers that facilitate the virus's entry into the bloodstream (a direct health effect), knowledge acquired in school about transmission risks (a direct schooling effect), and the human capital-marriage market pathway above.

After explaining these theoretical links, I then examine the empirical relationship

[3]Men 15-24 are 12 percent, and the remaining men and women under 50 are each around 30 percent.

[4]In contrast to soil-transmitted helminths such as hookworm, these schistosomes are transmitted through freshwater and infest the bladder rather than the intestines. The other form of schistosomiasis in Sub-Saharan Africa is intestinal (caused by *S. mansoni*), though it was substantially less common in Zimbabwe, as I discuss in Section 2.2.3.

between urogenital schistosomiasis and HIV prior to the deworming program. Combining rates of heavy schistosome infection—the main driver of morbidity—among students in 71 of Zimbabwe's districts with Demographic and Health Surveys (DHS) data from 2005 and 2010, I find strong pre-treatment correlations between heavy infection levels and HIV prevalence among the broader population. Importantly, the patterns in correlations by age group are consistent with the importance of age-disparate relationships in explaining the schistosomiasis-HIV connection: it existed among older men, young women, and older women but not among young men, who had not yet aged into the high-HIV transmission cycle mentioned above.

To test this relationship more rigorously, I exploit the quasi-experimental variation in urogenital schistosomiasis among school-age girls generated by deworming. As I describe in Section 2.3, rates of *S. haematobium* infection fell significantly after the first annual round and they remained low even prior to subsequent ones. Consequently, some districts experienced much greater reductions than others in rates of urogenital schistosomiasis for girls, so I use DHS data from the two pre-program years and 2015 in a difference-in-differences approach to compare pre- and post-treatment trends in outcomes among women aged 18 to 20 across (formerly) high- and low-schistosomiasis areas. I focus on these women because they were of reproductive age when deworming began in 2012 and much more likely to have been in school—and thus dewormed—given the steep decline in attendance after age 17. As placebo tests, I also compare trends for women aged 21 to 23 (because they met the first criterion but not the second) and men of the same age groups (because far fewer young men had HIV).[5]

I present the HIV results in Section 2.4. Three years after deworming began, HIV rates for young women in the age group more exposed to treatment fell by 4.3 percentage points (p.p.) more in formerly high-schistosomiasis districts, which was 47 percent of their pre-treatment mean. As expected, there was no effect detected for women in the mostly untreated age group, further suggesting that the reduction in urogenital schistosomiasis in school-aged girls explains this result. There was also no detectable impact on young men's HIV rates, conforming with the lack of a correlation between heavy schistosome infection and prevalence of the virus in this group.

In Section 2.5, I examine outcomes related to the three proposed channels linking urogenital schistosomiasis and HIV, and I find somewhat noisy but strikingly consistent evidence supporting the human capital-marriage market explanation. Following deworming, women ages 18 to 20 in formerly high-schistosomiasis areas were 7.1 p.p. (51

[5]It is therefore possible that null results for men would arise simply for this reason (i.e., floor effects). The direction of their HIV results and the evidence on channels should suggest whether it is the case.

percent) more likely to still be attending school. In contrast, I cannot reject null effects on school attendance for the slightly older women or men of the same age.[6] I then turn to marriage market outcomes and find similar patterns. After verifying that marriage rates did not change differentially, I find that age disparities between young women more exposed to treatment and their partners decreased by 2.0 years (30 percent) more in formerly high-schistosomiasis districts. In addition to this substantial decrease in an important HIV risk factor, I find that 3 measures of relationship power increase more for these young women: the ability to refuse sex (9.0 p.p., or 13 percent), condom use (2.6 p.p., or 426 percent), and say in one's own health care decisions (15 p.p., or 18 percent).[7]

Next, I perform a back-of-the-envelope quantification of the contribution of the likely reduction in genital ulcers to the HIV result for young women dewormed as girls. To do so, I take the approximate decrease in heavy schistosome infection and combine it with parameters from a simulated cost-benefit analysis by Ndeffo Mbah et al. (2013): the probability of having a genital ulcer if infected with *S. haematobium*, the rate of HIV transmission between serodiscordant couples, and how much that rate increases if the woman has a genital ulcer. It suggests that only 0.6 to 1.3 p.p. (15 to 30 percent) of the HIV result should be due to the direct health effect of urogenital schistosomiasis, implying that the human capital-marriage market channel is more important.

Lastly, I conduct a preliminary cost-benefit analysis of deworming as a method of HIV prevention program in Section 2.6. Taking the perspective of the Zimbabwean government—which may need to bear an increasing share of the costs of treating HIV as donor support declines—I use the size of the decrease in HIV and parameters from Ndeffo Mbah et al. (2013) to compute the expected present value of its costs and benefits. My calculations suggest that as a result of avoiding expenditures on lifelong ART and other health care arising from HIV infections, Zimbabwe's government will save nearly twice as much as was spent on deworming and increased secondary school attendance. Importantly, this approach does not consider the higher expected wages

[6] A treatment effect on young women's secondary school attendance but on not young men's matches the findings that Baird et al. (2016) reported in their 10-year follow-up of deworming in Kenya.

[7] I focus on these outcomes because of UN Population Fund Director Thoraya Obaid's criticism of HIV-prevention programs emphasizing abstinence, being faithful, and using condoms as failing to account for the realities facing women in developing countries:

> Abstinence is meaningless to women who are coerced into sex. Faithfulness offers little protection to wives whose husbands have several partners or were infected before marriage. And condoms require the cooperation of men. The social and economic empowerment of women is key. The epidemic won't [reverse without ensuring] ... women's right to sexual and reproductive health. (UNAIDS, 2004, p. 2)

or disability-adjusted life years (DALYs) that Miguel and Kremer (2004) considered, suggesting that there should be a substantial contribution of averted HIV infections to the cost-benefit analysis of deworming in high-prevalence countries in Sub-Saharan Africa.

Taken together, these results contribute to several strands of the economics and public health literatures. First, I show that improving the health of girls is a cost-effective method of preventing later HIV infection when they will be at exceptionally high risk of contracting the virus as young women. This paper thus adds to the analyses of Africa's devastating HIV epidemic that apply insights from economics one of the greatest public health crises of our time (Dupas, 2011; Robinson and Yeh, 2011; Oster, 2012; Björkman Nyqvist et al., 2018; Greenwood et al., 2019; Angelucci and Bennett, 2021).

More broadly, this paper furthers our understanding of the short- and long-run benefits of childhood health across countries and throughout history (e.g., Case et al., 2005; Maccini and Yang, 2009; Hoynes et al., 2016). It also shows that a health improvement later in childhood can still have important impacts for treated individuals as well as the fiscal health of governments. Therefore, it can potentially expand our view of the "better early than late" approach when it comes to such interventions (Almond and Currie, 2011; Currie and Almond, 2011; Gertler et al., 2014).

In particular, this paper adds a *quickly-realized* benefit in a novel domain to incorporate into justifications for childhood health interventions, especially deworming (Miguel and Kremer, 2004; Bleakley, 2007; Baird et al., 2016; Ozier, 2018; Hamory et al., 2021). Because discounting substantially reduces the present value of labor market returns a decade or more in the future, averting (at least 3 years of) ART can have an important impact on cost-benefit analyses, even if there is a lag between infection and treatment initiation.[8] The previously unexplored set of benefits in these results expands our view of the impacts of childhood health, raising the possibility that focusing exclusively on labor market returns understates the true value of interventions in this area.

Finally, this paper highlights the role of childhood health in marriage markets, especially in a non-Western context. The importance of human capital for marriage and fertility has largely been established by studying education (Aaronson et al., 2014; Duflo et al., 2015; Chiappori et al., 2018), but health is also a major component of human capital and can affect decisions regarding the latter (Bleakley and Lange, 2009; Rocha and

[8]There has been recent controversy over studies of deworming—see the discussion in Aiken et al. (2015), Davey et al. (2015), Hamory Hicks et al. (2015), and Hargreaves et al. (2015). The relevant Cochrane Review is skeptical of the body of evidence on its health effects—see the discussion in Taylor-Robinson et al. (2019) and Croke et al. (2022). But even when taking the approach of GiveWell (2017), which continues to recommend donating to deworming charities because its extremely low costs are outweighed in expectation by uncertain but potentially substantial benefits, the HIV effects found in this paper help to increase the magnitude of the benefit side of the ledger, if not its certainty.

Soares, 2010). My results suggest that improving childhood health can improve inputs into and outcomes of marriage market matching, which has substantial consequences for women's welfare in the developing world (Ashraf et al., 2020; Corno et al., 2020).

2.2 Urogenital Schistosomiasis and HIV in Zimbabwe

In this section, I first discuss the relevant information regarding morbidity from urogenital schistosomiasis. I then review the hypothesized links between its symptoms and HIV. Lastly, I present the positive correlations between *S. haematobium* infection intensity and HIV prevalence in Zimbabwe prior to the nationwide deworming program, and I discuss what they suggest about the channels underlying this relationship.

2.2.1 Urogenital Schistosomiasis Morbidity

Colley et al. (2014) provided a thorough overview of human schistosomiasis that I briefly summarize here. Adult male and female worms live for 3 to 10 years in the veins near the bladder of an infected person, where they mate and release fertilized eggs. These eggs either become trapped in the tissue of the urogenital system or are discharged in urine into the environment, where those reaching freshwater will hatch and undergo a process that allows them to infect other humans.[9]

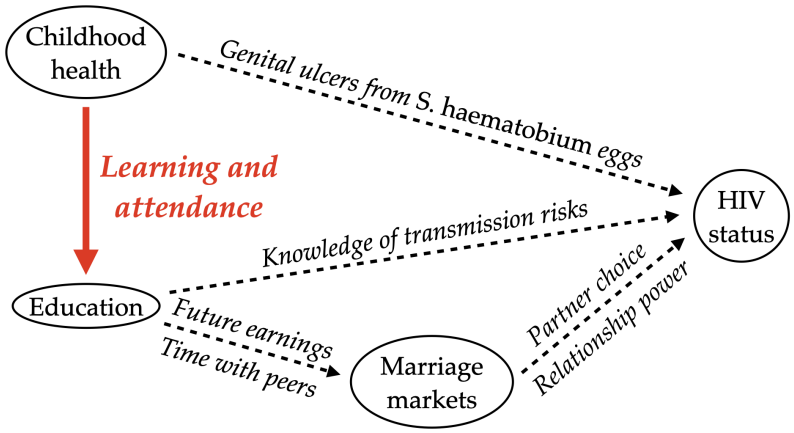
The cause of morbidity from schistosome infection is the immune response to eggs lodged in tissue. In particular, clusters of white blood cells surround trapped eggs and generate inflammation within the urogenital system.[10] Those infected for the first time at later ages than usual (most often adult travelers to endemic regions) can develop symptoms of acute schistosomiasis including fever and malaise lasting for 2 to 10 weeks.

Children born in endemic areas rarely experience this type of morbidity and instead suffer from chronic schistosomiasis. One of its most commonly discussed symptoms is anemia of inflammation—generally a mild to moderate anemia in which there are too few red blood cells but their size and hemoglobin content are normal (Weiss et al., 2019)—but whether it results from schistosomiasis alone or polyparasitic infections is not certain (Friedman et al., 2005). Chronic urogenital schistosomiasis can also cause organ-specific symptoms like blood in urine (haematuria, often mistaken for menses) and

[9]Specifically, the hatched eggs release swimming larvae that can infect certain snail species. These snails later release a different form of larvae that can penetrate the skin of humans coming into contact with the body of water. They then migrate to the target organs and mature into adult worms.

[10]This aspect of schistosomiasis is notably different from other helminthiases, in which morbidity is driven by worms consuming the host's blood and nutrition.

Figure 2.1: Linking Childhood Human Capital to HIV via Urogenital Schistosomiasis



Notes: Black text and dashed arrows denote channels examined in the existing literature. Red bold text and solid arrow denote the novel channel explored in this paper. See the text for references.

painful urination. In females, *S. haematobium* eggs can become lodged in the genital tract and result in lesions (i.e., genital ulcers).

2.2.2 Theoretical Links to HIV

There are several important ways in which these symptoms could affect HIV status through their impacts on the components of childhood human capital. Figure 2.1 summarizes the existing hypotheses regarding human capital and HIV, highlights the novel one examined in this paper, and specifies how urogenital schistosomiasis would factor into these channels. First, the public health literature focuses primarily on the disease's direct health effect: namely, ulcers caused by *S. haematobium* eggs lodged in the female genital tract that can facilitate the virus's entry into the bloodstream (Kjetland et al., 2006).[11]

Separately, economists and public health scholars have linked the education component of human capital to HIV through knowledge of how the virus spreads and schooling's impacts on sexual behaviors (Case and Paxson, 2013; Agüero and Bharadwaj, 2014; Behrman, 2015).[12] The idea behind the latter is that schooling heavily influences whom young women match with in marriage markets and their power in a relationship

[11]Urogenital schistosomiasis in men can also promote the spread of the virus: if they are co-infected with *S. haematobium* and HIV, their semen contains more virus-hosting cells and viral RNA as a result of their genital ulcers (Leutscher et al., 2005; Midzi et al., 2017).

[12]De Neve et al. (2015) showed that Botswana's 1996 expansion of secondary schooling reduced the risk of HIV infection as adults but the census data used did not permit an investigation of the underlying channels. Also related is Baird et al. (2012), who found that cash transfers conditional on school attendance reduced the odds of HIV infection for females aged 13 to 22 by changing sexual behaviors.

(Becker, 1991). Schooling may have such effects because of how it affects future labor market prospects (Peters and Siow, 2002), or simply because more time spent in school could increase the share of peers in the pool of potential matches.

But in spite of the prominence of the link between childhood health and education in studies from both the developing and developed world (e.g., Glewwe and Miguel, 2007; Currie, 2009), its role has not yet been explored in the context of HIV. Because better health in school-age years can increase learning and attendance (see Bleakley, 2010a, for a review), an increase in young women's educational attainment due to deworming could thus improve their chances of finding less-risky partners. And once in relationships, young women with more schooling could be better able to negotiate with their partners to reduce HIV exposure risks (UNAIDS, 2004).

2.2.3 Empirical Links to HIV

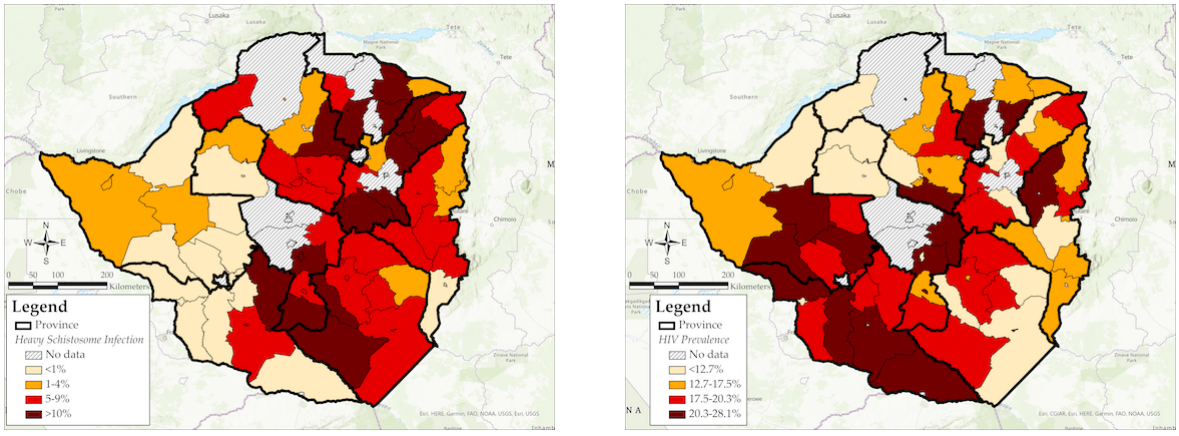
2.2.3.1 Data

Given the multiple channels through which urogenital schistosomiasis could affect HIV prevalence, I examine whether the intensities of these two diseases were in fact correlated in Zimbabwe prior to its deworming program. I use data on the former from Midzi et al. (2014), who reported results from the 2010-11 national helminthiasis survey conducted by the national government. Tests of over 13,000 students from nearly 300 schools across the country showed that rates of infection by either type of schistosome (above 22 percent) were over 4 times higher than rates for any soil-transmitted helminth (under 6 percent), with much greater rates of both helminthiases in rural areas. And of the two forms of schistosomiasis, urogenital (18 percent) was more than twice as common as intestinal (under 8 percent).

Nonetheless, heavy infection is generally the primary determinant of helminthiasis morbidity (Hotez et al., 2006), and it is especially true for urogenital schistosomiasis (Wiegand et al., 2021).[13] Because heavy infection rates for *S. haematobium* were nearly 20 times higher than for *S. mansoni* (5.5 percent versus 0.3 percent), I focus on urogential schistosomiasis as the disease most affected by deworming and use these rates as the relevant measure of the morbidity burden among school-age children. Figure 2.2a shows the variation in heavy *S. haematobium* infection in schools across districts (second-level administrative units) in 2010-11. The eastern half of the country had most of the urogenital schistosomiasis morbidity, but there was still significant variation within provinces.

[13]Heavy *S. haematobium* infection is defined as having visible haematuria (blood in urine) or at least 50 eggs per 10 milliliters of urine. A lower egg count constitutes a light infection (WHO, 2002).

Figure 2.2: Heavy Schistosome Infection and HIV Prevalence



(a) Heavy *S. haematobium* Infection, Students (b) HIV Prevalence, Ages 15-49

Notes: Maps show the respective disease intensities at the district level, with darker colors representing higher prevalence. Ranges of heavy schistosome infection rates among students are taken from Midzi et al. (2014) and correspond to low (15 districts), moderate (13), high (27), and highest morbidity (16). HIV prevalence is calculated from the 2005 and 2010 DHS surveys using blood test results from between 14 and 774 respondents in each district (median: 209). Levels of prevalence are grouped into quartiles.

I compare these data to HIV prevalence measured in the pre-deworming waves of the DHS (2005 and 2010; see Appendix Figure B.1 for a map of survey clusters by year). In these surveys, random subsets of respondents were offered anonymized HIV tests, and those that consented had their blood drawn.[14] Figure 2.2b shows districts' HIV positivity rates in pre-treatment years, which were highest in the south and east of the country.

2.2.3.2 Measuring Correlations

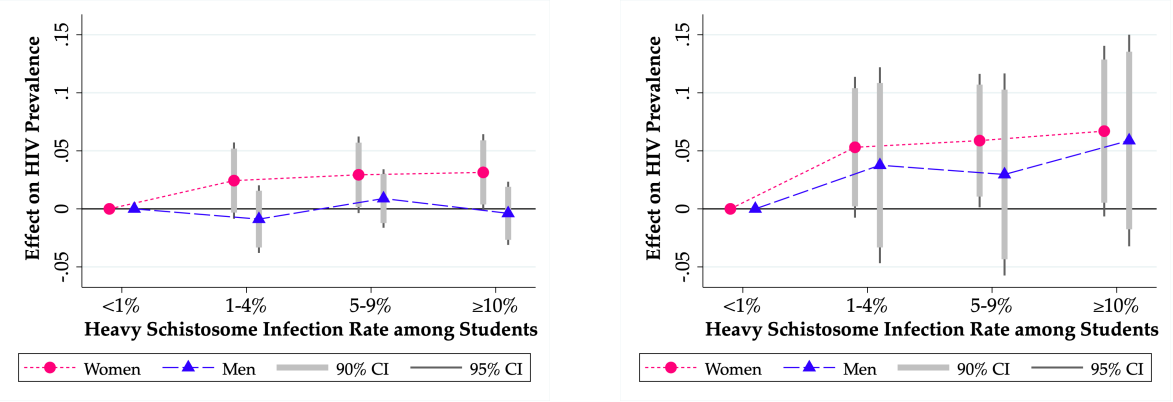
The visual similarity of these maps suggests a correlation between schistosomiasis morbidity and HIV prevalence. I examine this relationship formally by assigning DHS survey clusters their district's category of heavy schistosome infection and estimating

$$HIV_{i,c,t} = \alpha_{p(c)} + \gamma_t + \sum_{k=2}^{4} \tau_k \cdot SchistoCategory_{d(c)} + \mathbf{X}_i\beta + f(Lat_c, Lon_c) + \epsilon_{i,c,t}, \quad (2.2.1)$$

where $HIV_{i,c,t}$ indicates whether individual i in survey cluster c in year t is HIV positive, $\alpha_{p(c)}$ and γ_t are fixed effects for c's province and the year, $SchistoCategory_{d(c)}$ is c's district's category of heavy schistosome infection (corresponding to those in Figure 2.2a, with 4 being the highest), X_i are individual-level controls (age and age squared), $f(Lat_c, Lon_c)$ is a quadratic polynomial in c's latitude and longitude coordinates, and $\epsilon_{i,c,t}$ is the idiosyncratic error term.

[14] In pre-treatment years, 11 percent of women and 16 percent of men refused these tests.

Figure 2.3: Pre-Treatment Cross-Sectional Relationship between Schistosomiasis and HIV



(a) Ages 15-34

(b) Ages 35-54

Notes: Plots show the regression-adjusted relationships between districts' categories of heavy schistosome infection for students and HIV prevalence for each age group and sex in 2005 and 2010, with coefficients estimated relative to the lowest category (<1%). Categories are taken from Midzi et al. (2014). Regressions control for year and province fixed effects, age, age squared, and a quadratic polynomial in latitude and longitude. Standard errors are clustered by the 67 districts from which there are observations in 2005 and 2010. Regressions use 7,625 observations for females and 5,656 for males in (a) and 2,673 for females and 2,134 for males in (b).

The coefficients of interest are the τ_k, which measure the effect of a district being in the given category of heavy schistosome infection relative to being in the lowest one (<1%). These estimates come from comparing individuals across districts within provinces and years after adjusting for the individual- and cluster-level controls. For inference, I cluster standard errors by the 67 districts represented in the 2005 and 2010 DHS data.

2.2.3.3 Results

Due to the importance of age-disparate relationships for the spread of HIV in Southern Africa (see the discussion in Section 2.1), I split the data by sex and age group to study whether schistosomiasis intensity appears to interact with this transmission cycle. Figures 2.3a and 2.3b show the respective ecological relationships between schistosomiasis and HIV for these groups before the deworming program began. HIV prevalence increased with schistosomiasis morbidity for both younger and older women, but among men, there was only such a relationship for the older group. These patterns are what would be expected if poor childhood health contributed to the de Oliveira et al. (2017) HIV transmission cycle driven by age-disparate relationships, in which older men pass high infection rates to younger women who, as they age, pass them on to their male peers.

2.3 Empirical Strategy

While the theoretical and empirical links between urogenital schistosomiasis and HIV described above are clear, they do not imply a causal relationship or shed light on the underlying channels. To generate more rigorous evidence in this vein, I exploit the quasi-experimental variation in heavy *S. haematobium* infection among school-age children generated by deworming. Below, I briefly describe this campaign and then explain the difference-in-differences strategy I use to measure its impacts.

2.3.1 Success of Zimbabwe's Deworming Program

In September 2012, Zimbabwe conducted its first round of mass administration of the antihelmintic drug praziquantel in schools across the country. Five additional rounds followed in October 2013, January 2015, November 2015, November 2016, and November 2017. To measure the program's impact, Mduluza et al. (2020) selected a cohort of children across 35 sentinel schools with a range of pre-treatment *S. haematobium* prevalence rates to follow over time. The authors tested their urine for eggs and haematuria immediately prior to each round as well as 6 weeks later.

They found that the deworming program had rapid and sustained successes. After just one round of drug administration, the prevalence of *S. haematobium* infection in the cohort of tracked students fell from about one-third to around 1 percent. Some were reinfected between rounds: at the start of the second and third, rates had risen back to around 5 percent. But each time, drug administration significantly lowered infection rates again, and they never exceeded 2 percent at any subsequent point in the study.

2.3.2 Difference-in-Differences Strategy

The variation in pre-treatment heavy *S. haematobium* infection rates combined with the nationwide success of Zimbabwe's deworming program caused some districts to have substantially larger improvements in childhood health than others. I use this variation to identify its effects because deworming began at the same time in all districts. Specifically, I compare in each period those with high or the highest pre-program rates of heavy schistosome infection among schoolchildren (at least 5 percent) to those with low or moderate pre-program rates (below 5 percent), as heavy infection drives morbidity.[15] Forty-three districts were thus in the high category and 28 were in the low one.

[15]Zimbabwe's treatment guidelines were for (near-)universal drug administration among students where there was at least 5-percent heavy schistosome infection rates. Below that threshold, drug administration was to be much more targeted (Midzi et al., 2014).

To make these comparisons, I estimate the two-way fixed effects specification

$$y_{i,c,t} = \alpha_{d(c)} + \gamma_t + \sum_{k \in \{2005,\ 2015\}} \tau_k \cdot \left(HighSchisto_{d(c)} \cdot \mathbf{1}[t = k] \right) + \mathbf{X}_i \beta + f(Lat_c, Lon_c) + \epsilon_{i,c,t}, \tag{2.3.1}$$

which is similar to equation (2.2.1) except in the following respects: $\alpha_{d(c)}$ is a fixed effect for the district d in which a cluster lies, $HighSchisto_{d(c)}$ indicates whether a district had high or the highest pre-program schistosomiasis morbidity (i.e., heavy infection rate above 5 percent), and $\mathbf{1}[t = k]$ indicates whether an observation is from the given year k.

The coefficients of interest are the τ_k in 2005 and 2015, which measure the difference in an outcome between high- and low-schistosomiasis districts in the given year relative to the size of that difference in 2010 (the omitted year). An insignificant estimate in both the statistical and economic senses for τ_{2005} implies that pre-program trends across these districts evolved in parallel, and a significant τ_{2015} indicates that the improvements in young adults' health as children resulting in outcomes diverging. For inference, I cluster standard errors by district because treatment was assigned at that level.

As school-based helminth control began in 2012 and secondary school typically ends at age 17 in Zimbabwe, the only young adults who were both exposed to the program and of reproductive ages for all 3 years before the post-treatment DHS survey were between ages 18 and 20 in 2015 (i.e., ages 15 to 17 in 2012).[16] Therefore, I expect that the clearest effects should arise among young adults in this age range. As a placebo test, I also examine outcomes for young adults ages 21 to 23 (i.e., ages 18 to 21 in 2012) since they were close in age but almost entirely untreated and thus should not have directly benefitted from the deworming program.[17]

2.4 Results

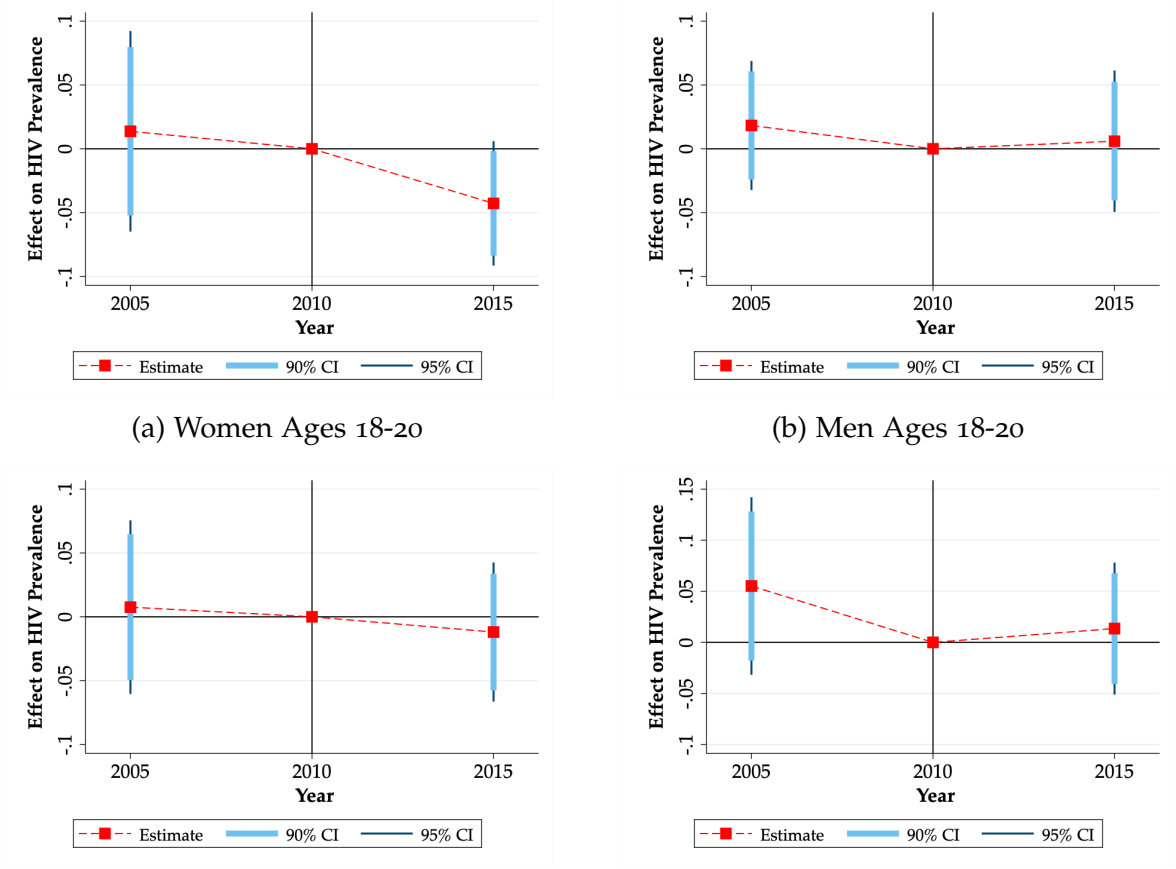
2.4.1 HIV Prevalence Declines for Young Women

I first examine the evolution of HIV prevalence among young adults exposed to school-based deworming. Figure 2.4a shows that for women ages 18 to 20, differences

[16] Prior to the program, just under half of girls and just under three-fifths of boys aged 17 attended school. Rates for 18-year olds are half of those attendance figures.

[17] Untreated groups still may have benefitted indirectly because nearby schoolchildren were no longer shedding *S. haematobium* eggs (e.g., Miguel and Kremer, 2004). But given these adults' ages when the program began, improved childhood health should not explain any benefit accruing to them, so there should not be evidence consistent with the education-related channels in Figure 2.1. Instead, any effect on HIV would likely arise from having fewer genital ulcers as adults (e.g., Kjetland et al., 2006).

Figure 2.4: HIV Prevalence among Young Adults



(a) Women Ages 18-20

(b) Men Ages 18-20

(c) Women Ages 21-23

(d) Men Ages 21-23

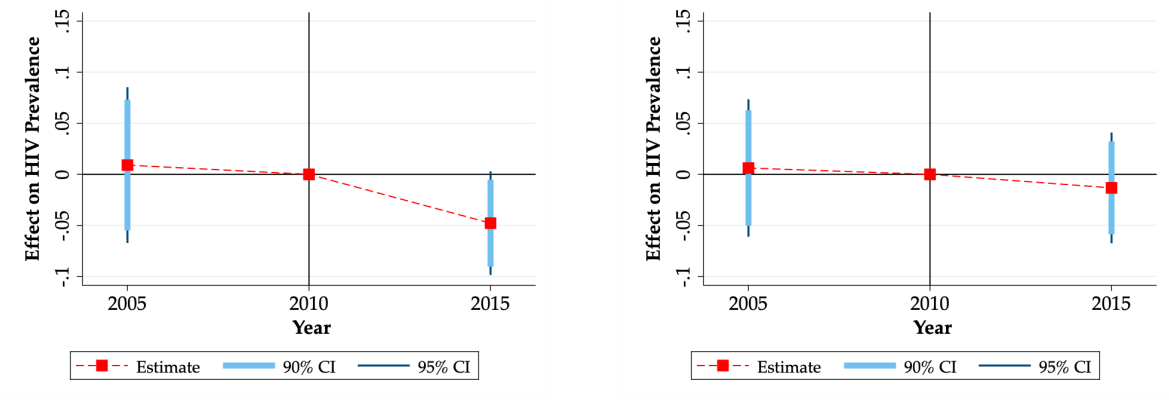
Notes: Plots show two-way fixed effect estimates for HIV prevalence among young adults of each sex. Regressions control for year and district fixed effects, age, age squared, and a quadratic polynomial in latitude and longitude. Standard errors are clustered by the 71 districts in the sample. Regressions use 2,010 observations in (a), 1,838 in (b), 1,868 in (c), and 1,210 in (d).

in seroprevalence between high- and low-schistosomiasis districts evolved in parallel prior to the program, but subsequently they diverged: HIV prevalence declined by 4.3 percentage points (p.p.) more in the former group. The estimated effect provides quasi-experimental evidence that improvements in girls' health leads to fewer HIV infections for them as young women, and its size—nearly half of their 9-percent pre-program prevalence—suggests that childhood health matters substantially for female HIV.[18]

In contrast, there is a negligible effect (0.6 p.p.) in Figure 2.4b of improving boys' health on their HIV status by ages 18 to 20. This null result is not surprising, as young

[18]When more than doubling the sample by expanding the age range to 15 to 20, a similar pattern is evident but precision is greater: pre-program trends are also parallel and the estimated treatment effect is 3.4 p.p. (SE 0.016), which is of a similar magnitude in percentage terms.

Figure 2.5: HIV Results after Controlling for Marital Status



(a) HIV Positive: Women Ages 18-20 (b) HIV Positive: Women Ages 21-23

Notes: Plots show two-way fixed effect estimates for HIV prevalence among young women of each age group after adding a single (i.e., never married) fixed effect to the covariates mentioned in Figure 2.4. Standard errors are clustered by the 71 districts in the sample. Regressions use 2,010 observations for ages 18 to 20 and 1,868 for ages 21 to 23.

men's low seroprevalence (2.8 percent prior to treatment) stems from their lack of sexual context with high-HIV risk groups. Taken together, these findings strengthen the story in Figure 2.3a: childhood morbidity affected young women's HIV but not young men's.

2.4.2 HIV Prevalence Declines Only for Female Cohorts Exposed to Treatment

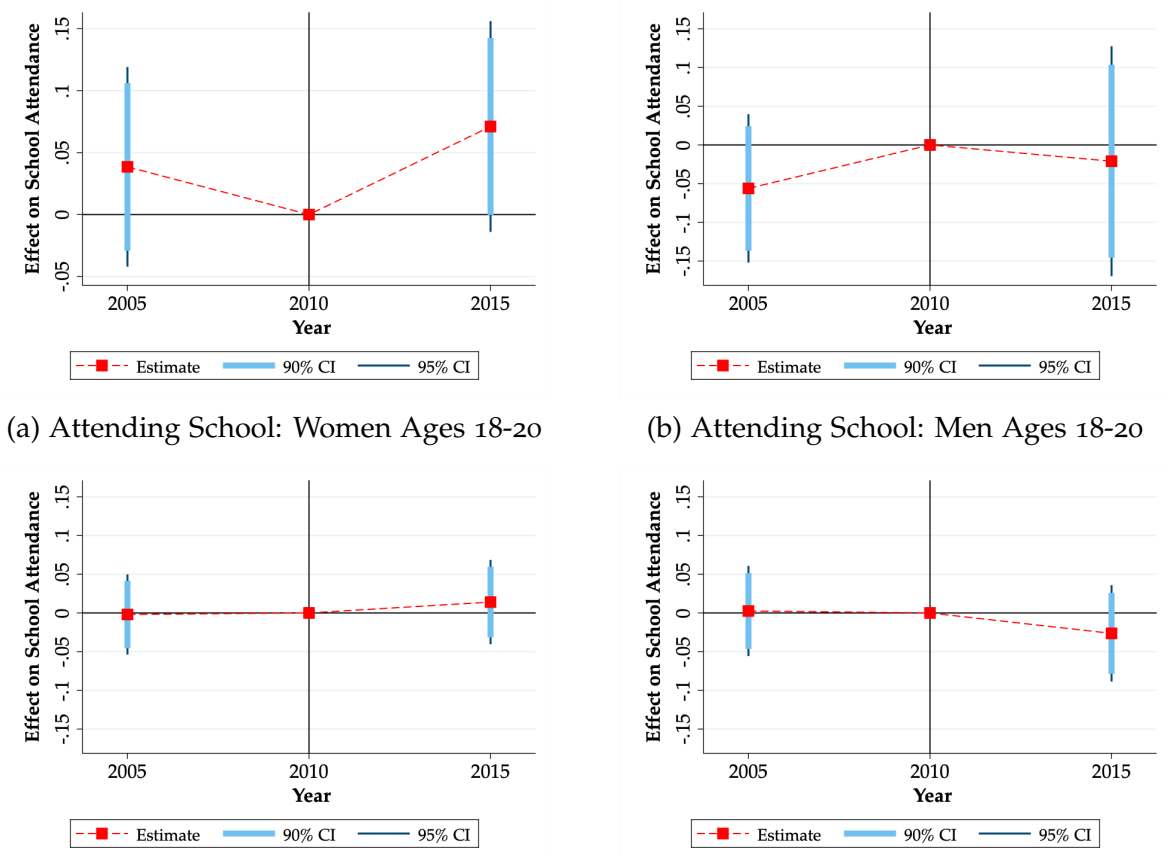
Next, in anticipation of null results, I study the trends in HIV prevalence for the slightly older cohort that was much less exposed to treatment (ages 21 to 23). Figures 2.4c and 2.4d are consistent with this prediction, as HIV prevalence declines by 1.2 p.p. more for young women in high-schistosomiasis districts (8 percent of their pre-program rates) and actually *increases* by a similar amount for their male counterparts (19 percent of their pre-program rates). In neither case are these figures distinguishable from 0, and once again the placebo treatment effect for men is smaller than their noisy estimated pre-trend.

2.4.3 Suggestive Evidence of a Role for Marriage and Dating

Though likely a "bad control" that is also an outcome of treatment (e.g., Angrist and Pischke, 2009), marital status may nonetheless help explain the above results due to its influence on the HIV risks facing young women. Therefore, some of the imprecision in Figure 2.4a could arise from failing to compare only within marital statuses.

After adding a single (i.e., never married) fixed effect to equation (2.3.1), Figure 2.5a shows that the effect is slightly larger and more precise than before (4.8 p.p., or just over

Figure 2.6: School Attendance among Young Adults



(a) Attending School: Women Ages 18-20

(b) Attending School: Men Ages 18-20

(c) Attending School: Women Ages 21-23

(d) Attending School: Men Ages 21-23

Notes: Plots show two-way fixed effect estimates for school attendance among young adults of each sex. Regressions control for year and district fixed effects, age, age squared, and a quadratic polynomial in latitude and longitude. Standard errors are clustered by the 71 districts in the sample. Regressions use 2,515 observations in (a), 2,492 in (b), 2,313 in (c), and 1,775 in (d).

half of pre-treatment HIV rates in high-schistosomiasis districts for this age group), while there is no substantive change in Figure 2.5b (-1.3 p.p.). With appropriate caveats, it suggests that marriage and dating play important roles in how girls' health shapes their HIV status as young women.

2.5 Channels

2.5.1 School Attendance Increases for Young Women

I then turn to a more rigorous examination of this and other channels that may explain the link between childhood health and women's HIV prevalence (see Figure 2.1). Figure 2.6a shows that school attendance for women ages 18 to 20 increased 7.1 p.p. more in

high-schistosomiasis districts, which is half of their pre-program rate. The effect appears to come from the younger part of this group remaining in high school: Appendix Figure B.2 shows that women ages 18 to 19 in high-morbidity districts became almost 8.6 p.p. more likely to be attending secondary school after treatment (just under half of their pre-program rate) but the estimate is imprecise. In contrast, there are no detectable effects on higher education attendance, though pre-program rates were very low (1.5 percent).

As with HIV prevalence, Figure 2.6b does not show any effect of treatment on young men's school attendance. This result could reflect childhood health having equal effects on young men's returns from entering the labor force and those from remaining in school, leading to ambiguous impacts on attendance (e.g., Bleakley, 2010a). However, because comparatively little is known about male rather than female genital schistosomiasis (e.g., Kayuni et al., 2019), it is also possible that the morbidity or stigma (or both) are less likely to keep boys and young men from learning and attending school.

2.5.2 School Attendance Increases Only for Female Cohorts Exposed to Treatment

I also study trends in attendance among young adults who were too old to have been exposed to treatment. Figures 2.6c and 2.6d show no detectable effects for women (1.4 p.p.) and men (1.9 p.p.) ages 21 to 23, though pre-program rates were understandably low for the former (3 percent) and the latter (8 percent).

2.5.3 Partner Age Gaps Decrease for Young Women

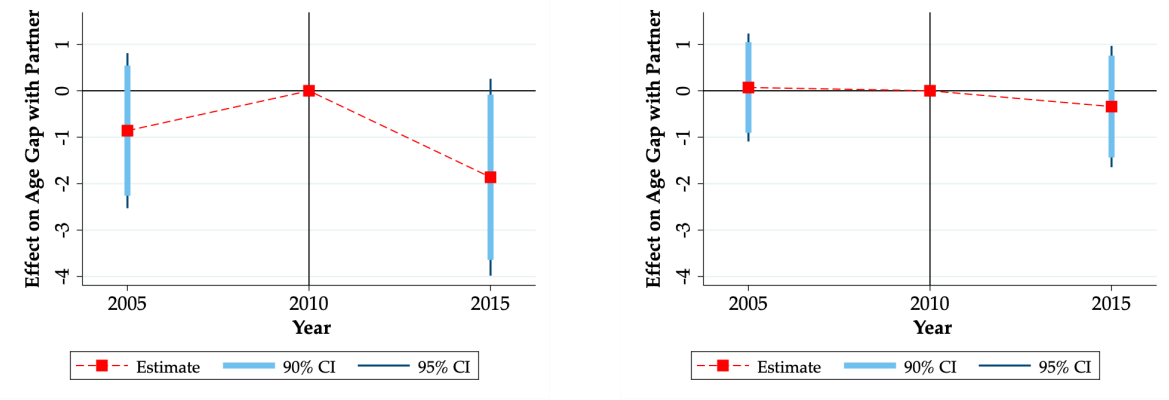
Next, I compare trends in whom young women match with in marriage and dating markets. After verifying in Appendix Figure B.3 that marriage rates for young women continue to evolve in parallel after treatment (potentially mitigating the econometric concerns in Section 2.4.3), I examine the age gap between young women and their partners given its major role in Southern Africa's HIV epidemic (e.g., de Oliveira et al., 2017).[19] I calculate it as his age minus hers and winsorize it at 90 percent due to outliers.[20]

Figure 2.7a shows that the age gaps between young women dewormed as girls and their partners decreases by an additional 2.0 years in high-schistosomiasis districts after treatment. This effect is nearly one-third of their mean pre-program partner age gap of 6.6 years, though its confidence intervals are wide. Conversely, in Figure 2.7b there is only a negligible decrease in partner age gaps for women ages 21 to 23 (-0.3 years).

[19]Half of women ages 18 to 20 and a quarter of women ages 21 to 23 were single prior to the program.

[20]The minimum pre-program age gap among all married women was -22 years, the 5th percentile was 0, the median was 5, the 95th percentile was 18, and the maximum was 77.

Figure 2.7: Partner Age Gaps among Young Adults Treated in Childhood



(a) Partner Age Gap: Women Ages 18-20 (b) Partner Age Gap: Women Ages 21-23

Notes: Plots show two-way fixed effect estimates for partner age gaps for young women of each age group. The outcome is positive if a male partner is older than a young woman, and it is winsorized at 90 percent. Regressions control for year and district fixed effects, age, age squared, and a quadratic polynomial in latitude and longitude. Standard errors are clustered by the 71 districts in the sample. Regressions use 1,020 observations for ages 18 to 20 and 1,427 for ages 21 to 23.

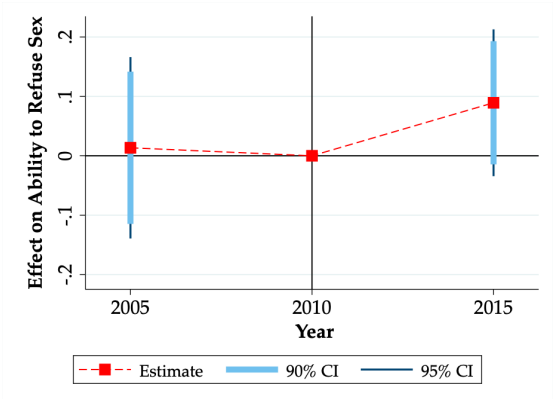
These results provide evidence that improving the health of girls leads them to match as young women with partners of lower HIV risk. Unfortunately, the data cannot speak to whether they did so with that goal in mind, because increased human capital reduced their need to rely economically on older men (e.g., Leclerc-Madlala, 2008), or simply as result of spending more time in school around male peers. Future research can shed light on this question.

2.5.4 Relationship Power and Safe Sex Increase for Young Women

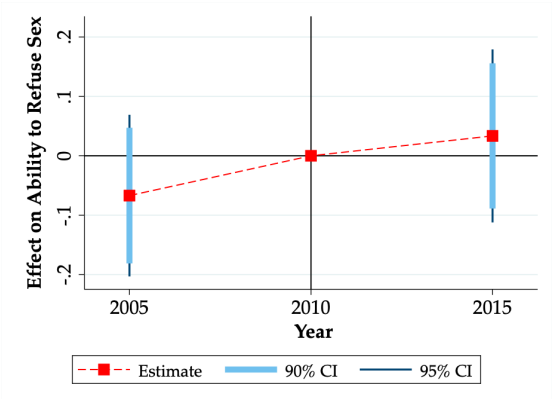
Finally, motivated by the shortcomings of some HIV prevention programs (see Section 2.1), I study whether improving the health of girls increases HIV-relevant measures of their relationship power as well as safe-sex behaviors. Figure 2.8a shows that the ability to refuse sex for partnered women ages 18 to 20 increased by 8.9 p.p. more in high-schistosomiasis districts following treatment, which was over 12 percent of pre-program rates, though the estimate is imprecise. But there were no comparable changes among cohorts less exposed to treatment (Figure 2.8b).

Similarly, in Figure 2.8c, condom use by partnered women ages 18 to 20 increased by 2.6 p.p. more after deworming in districts with high pre-program schistosomiasis morbidity. Though small in absolute terms, this effect size is substantial relative to usage rates of 0.6 percent prior to treatment. There was also a high degree of noise in these estimates and in Figure 2.8d, but taken together they suggest that deworming increased

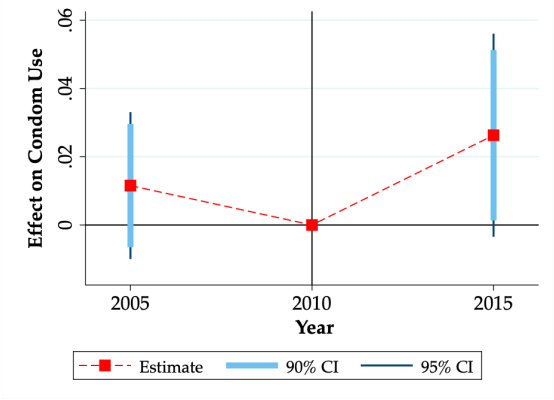
Figure 2.8: Relationship Power and Safe Sex among Young Women Treated in Childhood



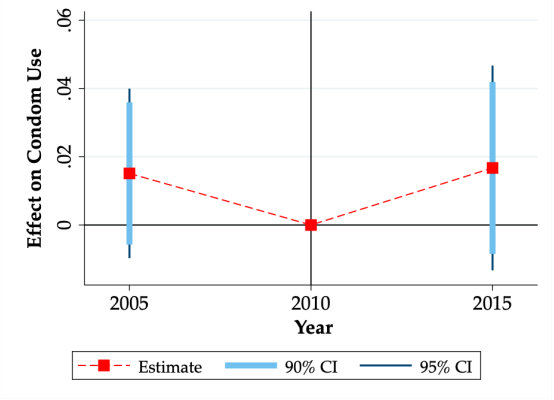
(a) Can Refuse Sex: Women Ages 18-20



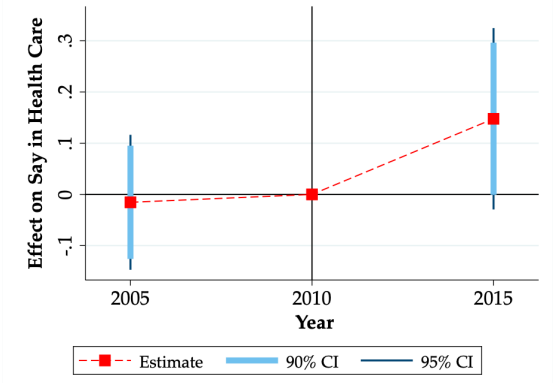
(b) Can Refuse Sex: Women Ages 21-23



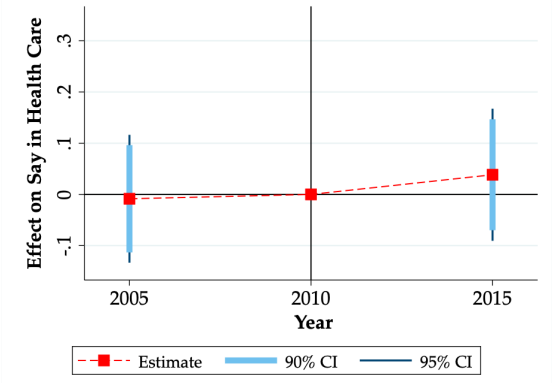
(c) Condom Use: Women Ages 18-20



(d) Condom Use: Women Ages 21-23



(e) Say in Own Health Care: Women Ages 18-20



(f) Say in Own Health Care: Women Ages 21-23

Notes: Plots show two-way fixed effect estimates for the respective outcomes. Regressions control for year and district fixed effects, age, age squared, and a quadratic polynomial in latitude and longitude. Standard errors are clustered by the 71 districts in the sample. Regressions use the following number of observations: (a) 1,021, (b) 1,429, (c) 2,278, (d) 2,075, (e) 1,008, and (f) 1,430.

the subsequent condom use of those exposed to treatment as girls.

I also examine trends in whether partnered young women were involved in or had control over decisions regarding their own health care. Figure 2.8e shows that this measure of HIV-relevant relationship power increased by 14.7 p.p. more for those ages 18 to 20 in high-schistosomiasis districts, which was 18 percent of pre-program rates. Conversely, partnered young women who were not exposed to treatment as girls experienced no discernible changes (3.8 p.p.) in who made their health care decisions (Figure 2.8f).

2.6 Toward a Cost-Benefit Analysis

The simplest assessment of deworming schoolchildren as an HIV prevention strategy focuses on the change in spending by the Zimbabwean government as a result of the program's impacts from 2012 to 2015. As I detail below, even from this myopic perspective, the program's benefits were nearly twice the costs it generated.

2.6.1 Benefits

For the government's finances, the benefit of averted HIV infections among young women was avoiding the lifelong costs of antiretroviral therapy and other health care arising from living with the virus. With a discount rate of 3%, the respective present values of these two expenditures in Zimbabwe convert to $3,750 and $868.75 in 2014 US dollars (Ndeffo Mbah et al., 2013). With a 4.3-p.p. decrease in HIV prevalence among women ages 18 to 20 due to the program, the discounted value of the expected reduction in the government's health outlays was $198.61.

2.6.2 Costs

In contrast, the present value of the various costs associated with the program was only $105.98. The smallest of these figures was the government's direct expenditures on schistosomiasis treatment. Converting what Ndeffo Mbah et al. (2013) reported to 2014 dollars, the recommended praziquantel dose cost $0.10 per student per year and the costs of administering the program were $0.26 per student per year. The first 3 years of the program thus cost $1.05 per student under a 3% discount rate.

Far larger are the costs arising from deworming increasing young women's secondary school attendance. UNESCO International Institute for Educational Planning et al. (2016) calculate that government spending per secondary school student was $328.20 in 2014. With a 7.1-p.p. increase in secondary school attendance rates for young women dewormed

as girls, the expected rise in school spending was $23.30 per student treated per year, or $67.89 per student treated in discounted terms.

Also important are health expenditures over the lifespan of the young women who avoided becoming HIV positive. In Zimbabwe, non-HIV health spending was $32.50 per person per year in 2014 dollars (Ndeffo Mbah et al., 2013) and life expectancy for 15-to-19-year-old women in 2015 was 50 years (WHO, 2020). Because HIV prevalence decreased by 4.3 p.p. among women ages 18 to 20 who were dewormed as girls, non-HIV health spending would increase by $37.04 per student treated in discounted terms.

2.7 Conclusion

These results show the importance of childhood health in reducing the spread of HIV in Sub-Saharan Africa, especially for the young women who are at great risk of contracting the virus. The effect appears at least in part to operate through the effect of human capital on their marriage and dating partners and the sexual behaviors they engage in, as it may shift them away from riskier choices in both areas. How exactly it occurs is unclear: it could be that they simply spend more time around similarly-aged men in the classroom or the result of improved labor market opportunities reducing their economic reliance on older men (or both).

Additionally, these results provide evidence of a novel benefit from controlling neglected tropical diseases, and helminthiases in particular. They show that along with its other important effects, a cheap intervention against a common childhood infection in a high-HIV prevalence country can also help combat one of the modern world's deadliest diseases. The fact that it can do so in an exceptionally cost-effective manner is also significant: donor funding to combat the global HIV pandemic has continued to decline, so the governments of some of the poorest countries in the world must bear a greater share of these costs moving forward.

CHAPTER III

Disease, Disparities, and Development: Evidence from Chagas Disease in Brazil (with Eduardo Montero)

3.1 Introduction

In the last two decades, the impact of disease on economic development has received much scholarly attention. While some studies argue for limited or even positive macroeconomic effects of disease in the modern era (e.g., Young, 2005; Acemoglu and Johnson, 2007; Ashraf et al., 2009), the large majority find substantial negative effects on a more micro level, with a focus on (anticipated) labor market returns (e.g., Miguel and Kremer, 2004; Bleakley, 2007). Importantly, they arise by affecting the children's human capital, whether through their ability to learn and attend school (Bleakley, 2010a) or expectations about the time horizon over which they will use it (e.g., Jayachandran and Lleras-Muney, 2009).

While this channel makes substantial contributions to long-run economic development by raising the adult incomes of those treated as children, there are two major downsides to relying solely on it to justify health interventions.[1] First, it takes nearly two decades for the first cohorts exposed to treatment for their entire childhoods to begin earning labor market returns from their increased human capital. As most of these returns are realized decades into their earnings trajectories, they are heavily discounted in cost-benefit analyses conducted when deciding whether to undertake an intervention. In addition, while schooling and labor income are easily measured, it is highly unlikely that they are the only development-relevant domains that health improvements affect. The result is that the benefit side of the ledger can be substantially understated.

[1]The main exception to the primary focus on childhood human capital has been the literature on HIV (e.g., Tompsett, 2020) because it primarily affects sexually-active adults, but it is still an important component of research on the virus's economic impacts (e.g., Baranov and Kohler, 2018).

This underestimation of the benefits of health interventions can be especially critical for neglected tropical diseases—those afflicting the poorest populations and causing them substantial morbidity—as donors may focus on quickly reducing mortality and developing-country governments may prioritize more immediate or certain payoffs.[2] Therefore, researchers can make substantial contributions to improving the lives and living standards of the world's poorest people by demonstrating that disease control has impacts in the short run as well as in the long run in areas relevant for economic development beyond individuals' labor market returns.[3]

We take a step in this direction by examining the effects of Brazil's efforts to eliminate the transmission of Chagas Disease (also known as American trypanosomiasis). As we discuss in Section 3.2, this parasitic neglected tropical disease almost exclusively afflicts the Americas south of the Rio Grande, and it affects both children and adults. Specifically, both groups can experience the acute phase's symptoms that last for weeks. And then around a decade or more later, a substantial share of those infected enter the chronic phase and develop cardiovascular problems, which affect the ability to work and are responsible for a substantial share of health problems in Latin America (Bocchi et al., 2009), as the disease affects around 6 million people in the region.

It is thus possible that in the short run, adults are unable to work due to the acute phase of Chagas Disease. In the long run, its chronic phase can also affect adult incomes and employment as well as hospitalizations and spending in Brazil's government-run health care system, which is the largest in the world and consumes 4 percent of GDP. Additionally, as poorer individuals are more likely to contract the disease—the vector lives in cracks in the roof and wall, which are more common in houses made of earthen materials—and given the close link between poverty and race in Brazil, the disease may also exacerbate the country's wide racial disparities.

We test these hypotheses by exploiting the geographic distribution of the main Chagas Disease vector prior to the control campaign. In Section 3.3, we describe the breakthrough that allowed for such a program, the 1975-83 nationwide entomological survey measuring the vector's presence at the municipal level, and the progress of insecticide spraying between 1984 and 1989. As we describe in Section 3.4, these data allow us to compare municipalities that were never infested with the vector and those that had it prior to spraying but were rid of it by 1989.

The idea behind our approach is that municipalities where the vector was eliminated

[2]In the case of soil-transmitted helminths, see Taylor-Robinson et al. (2019) for a systematic review that is cautious about the evidence for the benefits of mass deworming.

[3]To the extent that these benefits are public goods, they strengthen the case for collective action.

experienced a greater reduction in exposure to Chagas Disease than those where the vector was never found. We formalize this empirical strategy in Section 3.5 and use nationwide survey data from that period and a difference-in-differences framework to examine the short-run effects of the vector control campaign. As hypothesized, we find relatively rapid effects on employment rates for adults: moving from the 25th percentile of the probability of living in a treatment municipality to the 75th percentile results in a 1.3-percentage point (p.p., 1.9 percent) greater increase in the likelihood of being employed in post-spraying years. These effects are larger for non-white (1.9 p.p., or 2.6 percent) than white Brazilians (1.1 p.p., or 1.5 percent), suggesting that it raised the speed of racial convergence in the short run.

In Section 3.6, we turn to the long-run labor market results for adults treated as children using data from the 2010 census of Brazil and the same empirical strategy. When pooling racial groups, we find positive but imprecise increases in incomes. However, when we separate the sample into white and non-white adults, we find substantially larger effects for the latter (1.5-percent increase in incomes versus 0.2 percent). We examine educational attainment as a channel for these results but do not find evidence that it can explain a large share of the effect, suggesting that the effects of chronic Chagas Disease may play an important role in this reduction of racial disparities.

If reducing the occurrence of the disease's chronic phase contributed to the increase in incomes, it may also have resulted in substantial public finance effects through its impact on Brazil's government-run health care system. In Section 3.7, we examine hospitalizations, person-days spent in the hospital, and spending on hospital care covered by this system by modifying our differences-in-differences strategy to also compare the above outcomes due to circulatory system diseases against those due to all other causes (i.e., a triple-difference framework).

We find approximately 10-percent greater reductions in hospitalizations and person-days spent in the hospital due to circulatory diseases—which account for around 10 percent of these outcomes covered by Brazil's public health care system. Importantly, these declines started 10 years after vector elimination, which is about the point when the chronic phase symptoms of Chagas Disease would have begun to appear. These decreases appear to have led to a 6-percent greater decrease in spending on hospital care due to cardiovascular causes beginning at the same time, though our estimate is somewhat imprecise. Nonetheless, it is an economically significant magnitude given that circulatory diseases account for 20 percent of public hospital care spending.

Taken together, these results imply that controlling Chagas Disease had important benefits for a developing economy both in the short run (adult employment) and in

addition to individuals' labor market returns in the long run (public health care spending). They also suggest that, in a multi-racial country in which race and poverty are closely linked, combating diseases that primarily affect the poorest citizens can contribute to reductions in disparities between these groups. As such, we provide novel and important evidence of the benefits of controlling neglected tropical diseases in the developing world that should impact donors' and policymakers' decisions to do so.

3.2 Overview of Chagas Disease

3.2.1 Causal Agent and Vectors

The parasite *Trypanosoma cruzi* causes Chagas Disease. Around 90 percent of those infected contracted it from infected blood-sucking triatomine bugs, which live in cracks in roofs and walls and emerge at night to take blood meals from sleeping humans.[4] In Brazil, the most important vector species is *Triatoma infestans*, responsible for 80 percent of all transmission (Schofield and Dias, 1999).[5] *T. infestans* became domesticated and spread via rural settlements after Brazilians began clearing forests for agriculture and ranching in the south and southeast of the country in the late nineteenth century (Schofield, 1988).

3.2.2 Phases

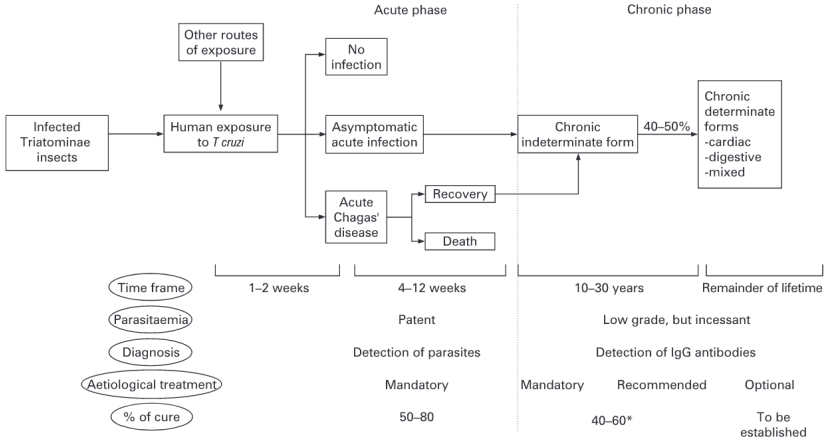
There are two phases of Chagas Disease. Figure 3.1 shows the progression of the disease from exposure through the rest of the patient's life. The acute stage begins after 1 to 2 weeks of incubation, lasts 4 to 12 weeks, and has non-specific symptoms like malaise and fever (Rassi et al., 2009). Children become more seriously ill in this phase than young adults, with as many as 10 percent of them dying from it, but more than 40 percent of those infected progress through this stage mostly symptom-free (Khan, 2011).

Individuals then enter the chronic phase. For 10 to 30 years, they experience no symptoms and the majority will remain in a chronic indeterminate stage in which they never develop any lesions (Rassi et al., 2000). However, the other portion progresses to the chronic determinate phase. Most commonly, this group develops cardiac complications such as the heart muscle becoming degraded and being replaced by fibrous tissue. Such cardiomyopathy is the cause of most of the morbidity and mortality from Chagas Disease (Nunes et al., 2018).

[4]The other 10 percent of transmission occurs through blood transfusions and the placenta.

[5]The other main Brazilian vector is *T. brasiliensis*, which accounts for 10 percent of all transmission. In Central and northern South America, the main vectors are *Rhodnius prolixus* and *T. dimidiata*.

Figure 3.1: Phases of Chagas Disease



Notes: Diagram taken from Rassi et al. (2009, p. 527).

3.2.3 Economic Consequences of Chagas Disease

Both stages have features that might affect economic development. Experiencing the acute phase in childhood should reduce lifetime income by affecting the knowledge component of adult human capital. Because of their symptoms, children would spend 1 to 3 months less able to focus in class or absent from school. In that case, they may fall behind their peers and repeat grades or decide to drop out entirely.[6] It is also possible for an adult to experience the acute phase's symptoms if they become infected, in which case they may struggle to work.

Entering the chronic determinate state of Chagas Disease in adulthood would have substantial negative impacts on developing countries as well. For individuals, years of shortness of breath, fatigue, and dizziness resulting from a heart with difficulty pumping blood should lead to reduced output at work, absenteeism, and even an inability to remain employed.

Therefore, Chagas Disease should also have broader consequences for developing countries. A universal public health insurance program like Brazil's *Sistema Único de Saúde* (SUS, which consumes around 4 percent of GDP) would have to spend more heavily on hospitalizations and doctor's visits for those suffering from acute and chronic symptoms. As the disease is most likely to be transmitted to humans living in houses with cracks in the walls and roof, it also disproportionately affects those who were already poor. In the

[6]In contrast to other parasitic infections (e.g., helminthiases), the medical literature on acute Chagas Disease does not mention sequelae like anemia and stunting that affect childhood development. But in a letter to the editor of *The Lancet*, Schofield (1981) estimated that individuals in the average house with *T. infestans* lose 2 to 3 ml of blood per day to the bugs, which might lead to anemia in children with inadequate diets or other parasitic infections. Subsequent studies did not investigate this claim, however.

context of a strong correlation between race and poverty, the disease may exacerbate gaps between white and non-white Brazilians.

3.3 Brazil's Vector Control Program

Much of this section summarizes Section 4 of Schofield and Dias (1999).

3.3.1 Breakthrough of Pyrethroid Insecticides

The post-World War II campaigns against malaria used organochlorine insecticides like DDT that were ineffective against Chagas Disease vectors. Several trials found that γ-benzene hexachloride (BHC) was effective if sprayed on the walls and roofs of triatomine-infested houses in high doses. In the 1960s, São Paulo's vector control superintendency began a program using BHC to effectively eliminate *T. infestans* from the state. However, São Paulo was the only state with the resources to implement such an intensive program.

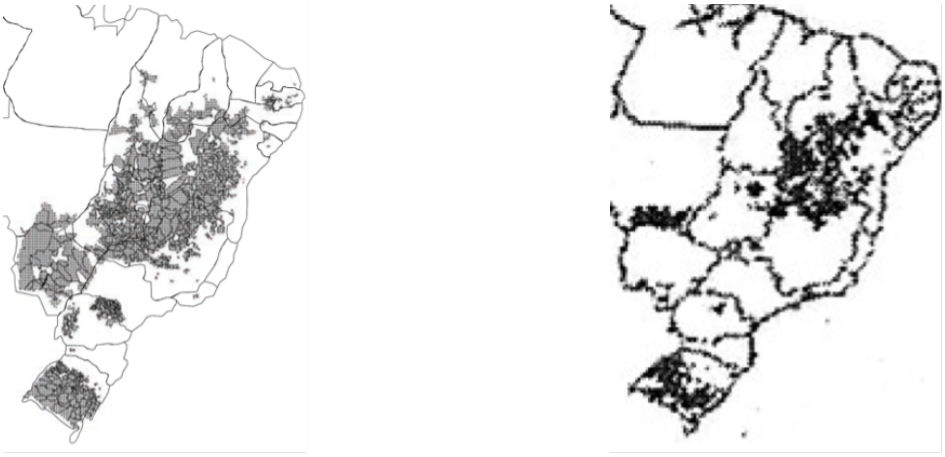
The second generation of pyrethroid insecticides became available in the 1970s and studies by the end of the decade showed their effectiveness against triatomine bugs. Importantly, they were effective when sprayed less frequently and at low doses, which made them more cost-effective than BHC in spite of a higher price per kilogram. They were also easy to apply and did not have unpleasant odors.

3.3.2 National Surveys and (an Interrupted) Vector Control Campaign

In 1975, the Brazilian government started a national campaign against Chagas Disease. It began because of, among other factors, the "development of suitable vector control methods, ... and continuous campaigning by scientists including demonstrations that vector control was feasible" (Dias, 1987, p. 338). The first stage consisted of serological and entomological surveys through the early 1980s. They found a national rural *T. cruzi* prevalence rate of 4.2 percent—including 8.8 percent rural prevalence in the heavily-infested states of Minas Gerais and Rio Grande do Sul—and vectors present in 36 percent of Brazil's territory (Dias, 1987). Figure 3.2a shows a map of municipalities that had *T. infestans* in dwellings from 1975 to 1983.

After the surveys concluded, thousands of sprayers visited millions of homes across the endemic region. However, the program came to a halt in 1986 due to the arrival of the *Aedes aegypti* mosquito in coastal areas and resulting outbreaks of dengue fever. Due to political pressure, the public health campaign superintendency diverted 40 percent of its

Figure 3.2: Progress of Vector Control Campaign, 1975-89



(a) *T. infestans*, 1975-83 (b) *T. infestans*, 1989

Notes: Figures taken from Silveira (2011) and Coura and Dias (2009) show municipalities with *T. infestans* in each period.

personnel to dengue control, leading to a reemergence of triatomine in recently-sprayed municipalities (Dias, 1987; Schofield and Dias, 1999). But vector control resumed in 1989, and the reduction in *T. infestans* in that year compared to 1975-83 is clear in Figure 3.2b.

To show the health impacts of (the interruption in) spraying, Figure 3.3 presents a graph of 1984-97 hospital admissions due to (acute) Chagas Disease.[7] They declined at the beginning of the campaign through 1986 and then increased with the interruption of control efforts lasting until 1989. Subsequently, admissions declined again as the program resumed, and by 2006, the Pan American Health Organization certified Brazil as having interrupted transmission of Chagas Disease through *T. infestans* in every state.

3.4 Data and Treatment Definition

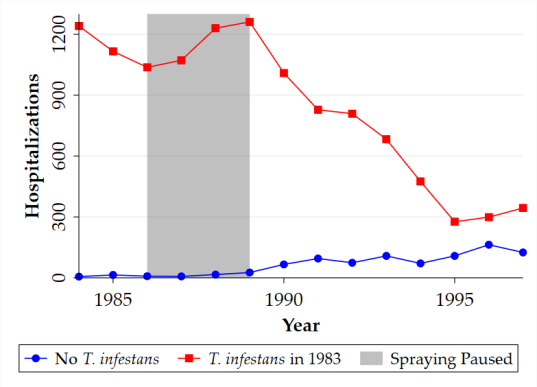
3.4.1 Data

Given the discussion of the potential effects of Chagas Disease in Section 3.2.3, we focus on both the short- and long-run economic impacts of the health improvement induced by control of triatomine. To study the former, we use microdata from the National Household Sample Survey (PNAD) for 11 of the 18 years from 1982 to 1999.[8] These surveys contain data on respondents' state of residence, household characteristics and sociodemographic information, and schooling and labor market outcomes. While

[7]Because of the non-specific symptoms of the acute phase, these numbers are almost certainly an undercount. Nonetheless, they are helpful in verifying the effects of the vector control campaign.

[8]The missing years are 1983, 1984, 1985, 1988, 1991, 1994, and 1996.

Figure 3.3: Hospitalizations for Acute Chagas Disease, 1984-97



Notes: Graph shows the evolution of hospitalizations for acute Chagas Disease (ICD-9 code 086) in states without (blue circles) and with (red squares) *T. infestans* in 1983. Data are from DATASUS. The shaded years (1986-89) denote the interruption in vector control.

the geographic resolution of these data are coarse and several years immediately around the start of spraying are missing, their near-annual frequency allows us to examine how quickly outcomes changed in states of residence after spraying began.

We also use IPUMS microdata from the 1980 and 2010 censuses of Brazil (Minnesota Population Center, 2020) to examine the long-run effects of Chagas Disease vector control. These 25- and 10-percent samples include information on a respondent's consistent 1980-2010 municipality of residence and consistent 1960-2010 state of birth, as well as an indicator for whether a respondent was born in their municipality of residence. We use the 1980 data to construct our measure of exposure to the treatment when only a respondent's state of residence or birth is known, as we describe in the next section. The census also contains more detailed household, sociodemographic, schooling, and labor market data than the PNAD datasets, which we exploit in the 2010 data.

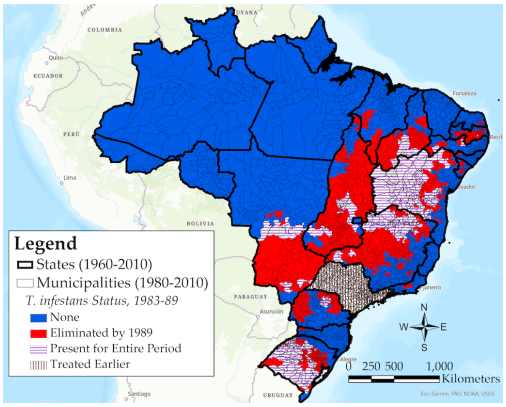
3.4.2 Treatment Definition

We define control municipalities as those that were never treated with spraying—i.e., they did not have *T. infestans* in 1975-83 and did not require vector control to achieve this status.[9] Our treatment group consists of those with *T. infestans* present in 1975-83 but eliminated by 1989 as a result of the spraying that began in 1984.[10] We use these years as endpoints because it ensures that our 2010 census sample has birth cohorts that were had spent their entire childhoods living without *T. infestans*. Because we cannot determine in

[9]The latter condition leads us to exclude the state of São Paulo, as its prior vector control campaign means it is not part of the never-treated group but rather an earlier-treated one (Goodman-Bacon, 2021).

[10]We also exclude not-yet-treated municipalities that had triatomine bugs in 1975-83 and 1989. Note that this group also includes municipalities where the vector to returned during the 1986-89 pause in spraying.

Figure 3.4: Municipalities by 1983-89 *T. infestans* Status



Notes: Map shows control municipalities (never had *T. infestans*) in blue and treatment municipalities (*T. infestans* eliminated by 1989) in red. Municipalities that still had *T. infestans* in 1989 (white with horizontal purple lines) or were treated prior to this period (white with vertical brown lines) are omitted from our sample. Data are from Silveira (2011) and Coura and Dias (2009) (see Figures 3.2a and 3.2b). The underlying shapefile of consistent 1980-2010 municipalities and 1960-2010 states is from IPUMS.

Table 3.1: Summary Statistics, 1980

	Municipalities by 1983-89 *T. infestans* Status			
	None	Eliminated	Present	Treated Earlier
Panel A. Demography				
Age	23.748	23.717	23.722	25.054
	(34.251)	(31.166)	(33.434)	(27.835)
Female	0.506	0.501	0.500	0.501
White	0.457	0.587	0.541	0.749
Brown	0.477	0.355	0.396	0.185
Black	0.065	0.056	0.059	0.046
Asian	0.002	0.003	0.004	0.020
Panel B. Labor Markets				
Log Monthly Income	4.258	4.261	4.002	5.374
	(4.361)	(4.343)	(4.283)	(4.532)
Employed (if in LF)	0.975	0.980	0.978	0.983
Panel C. Human Capital				
Years of Schooling	4.490	4.305	4.047	5.210
	(4.774)	(4.643)	(4.675)	(4.523)
Observations	15,328,696	4,228,428	1,914,177	5,686,026
Municipalities	1,156	402	173	309

Notes: Table lists means for variables of interest in the 1980 census and standard deviations in parentheses under the means for continuous variables. Categories correspond to Figure 3.4. Data are from IPUMS.

which year a municipality became free of the vector, we assume that all were treated in 1984 to be conservative.

Figure 3.4 shows control municipalities in blue, treatment ones in red, and those we exclude in white with either horizontal purple lines (not yet treated) or vertical brown ones (treated earlier). Table 3.1 shows summary statistics for demographic, labor market, and human capital variables among all individuals in these groups in 1980, 4 years before the start of spraying. Municipalities in our treatment group were more white and had slightly less schooling than those in our control group, though incomes were equal.

3.5 Short-Run Effects on Schooling and Labor Market Outcomes

We first examine whether reducing exposure to acute Chagas Disease affected a child's ability to attend school (ages 8 to 18) and an adult's ability to be employed (ages 35 to 50), as the 1 to 3 months of symptoms could interfere with both of these outcomes. Our results show that while spraying had positive but imprecise effects on school attendance, it noticeably increased adult employment within a few years, implying that policymakers did not need to wait decades for there to be economically meaningful returns to reducing Chagas Disease transmission. When examining results by racial group, we estimate that attendance and employment increased more (and more precisely) for non-white Brazilians, which would have helped to speed convergence in the country's racial disparities.

3.5.1 Empirical Strategy

Our evidence comes from comparing individuals of interest in each year of the PNAD (1982-2015) across states with varying levels of pre-treatment *T. infestans* presence, motivated by the idea that vector control induced greater improvements in health where there was more exposure to Chagas Disease prior to spraying. Our baseline estimating equation is the dynamic two-way fixed effects model

$$y_{i,s,t} = \alpha_s + \gamma_t + \sum_{k \neq 1986} \tau_k \cdot \left(\mathbb{P}[Treat]_s \cdot \mathbb{1}[t = k] \right) + \mathbf{X}_i \beta + \epsilon_{i,s,t}, \tag{3.5.1}$$

where $y_{i,s,t}$ is an outcome of interest for individual i living in state s in year t, α_s and γ_t are state and year fixed effects, $\mathbb{P}[Treat]_s$ is the probability in the 1980 census that an individual in a state of that sex and race resides in a treatment municipality, $\mathbb{1}[t = k]$ indicates whether an observation is from the given year k, $\mathbf{X}_i$ is a vector of individual-level covariates (age, age squared, and fixed effects for female sex and Asian, Black, and Brown racial categories), and $\epsilon_{i,s,t}$ is the idiosyncratic error term. Because we cluster standard

errors by the "small" number of states in our sample (24), we follow Cameron et al. (2008) and use the wild cluster bootstrap to generate confidence sets.

The coefficients of interest are the τ_k, which measure the difference in an outcome in a given year as the probability of residing in a treatment municipality goes from 0 to 1, relative to the size of that difference in outcomes in 1986.[11] Importantly, this choice biases estimates toward zero if the sign of the 1982 coefficient is the opposite of the treatment effect's, assuming that the size of the difference in outcomes in hypothetical 1983 data would have been between those in 1982 and 1986.

Estimating an effect for 1982 also permits a qualified assessment of pre-treatment trends, as a hypothetical 1983 coefficient should not have been "too different" from the 1982 estimate. Economically and statistically insignificant magnitudes for the pre-treatment year suggest that outcomes evolved in parallel prior to vector control, and significant estimates in years after spraying began indicate that outcomes diverged as a result of the decline in acute Chagas Disease.

Because nearly one-third of observations were from states with no treatment municipalities and the maximum probability of living in one was 0.740, the magnitude of these estimates overstates any policy-relevant parameter. Therefore, when discussing the results, we frame them in the context of moving from the 25th percentile of this probability to the 75th, which is an increase from 0 to 24.6 percentage points (p.p.)—i.e., this shift would have an impact of just under one-quarter of the estimated magnitudes.

Along with equation (3.5.1), we also estimate the static two-way fixed effects model

$$y_{i,s,t} = \alpha_s + \gamma_t + \tau \cdot \left(\mathbb{P}[Treat]_s \cdot \mathbb{1}[t > 1986]\right) + \mathbf{X}_i\beta + \epsilon_{i,s,t}, \tag{3.5.2}$$

which is identical to the previous equation except that the probability of residing in a treatment municipality is interacted with a single variable $\mathbb{1}[t > 1986]$, an indicator for whether a year is after 1986. By pooling the post-treatment years, the estimate of τ is likely to be more precise, though it comes at the cost of imposing a single value across all years in this period. We view these strategies as complementary because they both have an important drawback but each helps to address the concern regarding the other.

As with the dynamic estimates, the static estimate will be biased toward zero if the 1982 coefficient and the treatment effect have opposite signs. However, this bias will be less severe than in the dynamic estimates in this case because the static coefficient is estimated against the average of the difference in outcomes in 1982 and 1986. The result

[11] The final entirely pre-treatment year was actually 1983, but there are no PNAD data available for that year. We chose to make 1986 the reference year because it was only 2 years after spraying began and spraying paused in that year, making it only "lightly" treated.

is that the static coefficient will lie closer to the largest dynamic estimates (in absolute magnitude) instead of falling in the middle of the range.

Finally, to examine the potential for health improvements to reduce racial inequality in the short run, we estimate equations (3.5.1) and (3.5.2) separately for white and non-white individuals. Rather than conducting statistical tests for another difference in trends across racial groups (i.e., a triple-differences framework) given the potential for imprecision, we simply compare the respective estimates' magnitudes patterns and discuss implications.

3.5.2 School Attendance

We first examine the effects of Chagas Disease on school attendance in the years around the start of vector control. Figure 3.5a plots the estimates from equations (3.5.1) and (3.5.2) for school attendance among children ages 8 to 18. The patterns in the dynamic estimates suggest a positive but imprecise impact of reducing acute Chagas Disease exposure on attendance. However, it is important to note that they likely underestimate the true magnitudes given the opposite signs of the 1982 coefficient and the treatment effect.

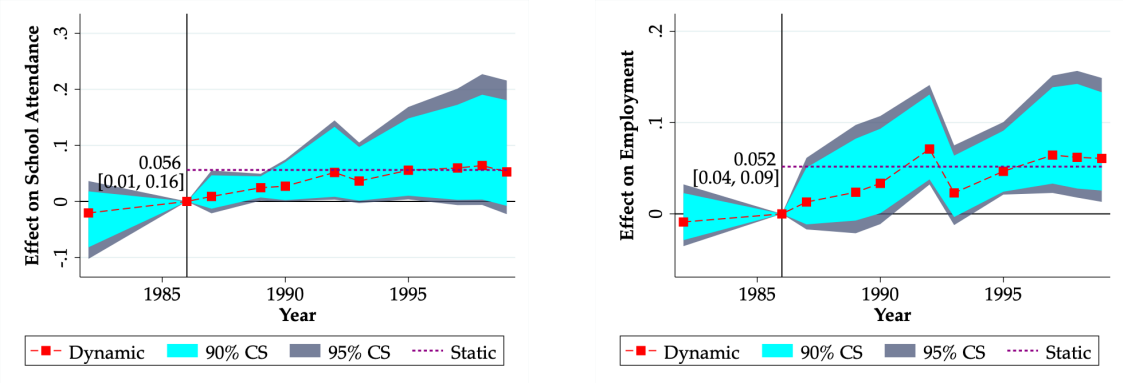
The static strategy recovers a coefficient of similar magnitude to the later years' dynamic estimates but with greater precision. The implied effect of moving from a 0 to one-quarter probability of living in a treatment municipality is 1.4 p.p. (1.9 percent of the pre-treatment mean) and it is significant at 10 percent. This estimate is likely closer to the true value than the dynamic estimates because it is calculated against the average of the 1982 and 1986 differences.

3.5.3 Adult Employment

Next, we study the short-run effects on employment (having worked in the past week) for adults ages 35 to 50 and plot the dynamic and static estimates in Figure 3.5b. There is little evidence of differential trends prior to the start of spraying, but in contrast to the attendance results, the divergence in employment trends occurs more rapidly and is more precisely estimated. By 1990, the implied dynamic estimates remain around 1.3 p.p. (1.9 percent), which equals the magnitude of the implied static estimate.

This result is particularly important because it shows that controlling Chagas Disease has relatively rapid effects on *adults already in the labor force*. While increasing the human capital of children makes large contributions to economic development, doing so yields benefits that are only realized after they enter the labor force a decade later and over decades of their working lives. Instead, triatomine control appears to have induced some unemployed adults to (re-)enter employment, which could have quickly improved

Figure 3.5: Short-Run Effects on School Attendance and Adult Employment



(a) Attending School, Ages 8-18

(b) Employed, Ages 35-50

Notes: Graphs plot dynamic (red squares) and static (dotted purple line) two-way fixed effects estimates for the respective outcomes with 90-percent (light blue) and 95-percent (dark blue) wild cluster bootstrap confidence sets for dynamic estimates. Static estimate magnitudes are next to the dotted purple lines with 90-percent wild cluster bootstrap confidence sets below them in brackets. Data are from the 1982-1999 PNAD surveys. Regressions control for fixed effects for state, year, female sex, and racial category (Asian, Black, and Brown) as well as age and age squared. Standard errors are clustered by the 24 consistent 1960-2010 states. Regressions use 836,331 observations for children and 1,198,480 observations for adults. In pre-treatment years, 74.7 percent of children attended school and 67.0 percent of adults were employed.

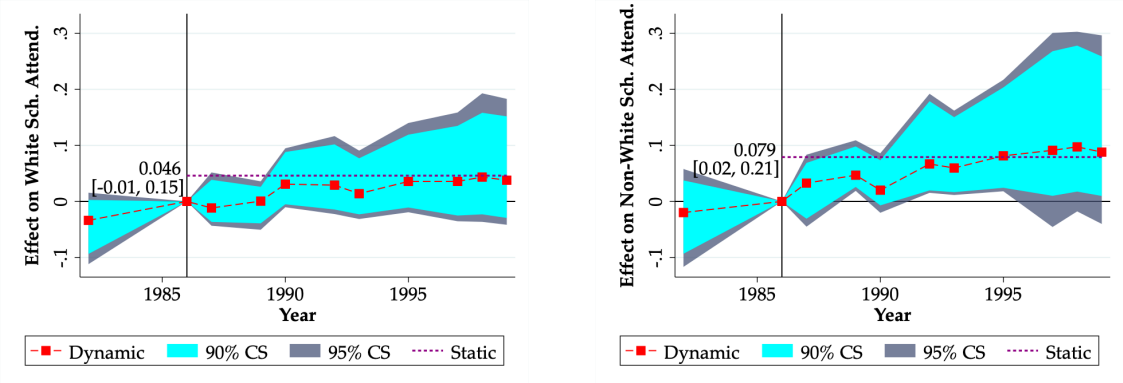
standards of living for them and their families.

3.5.4 Reducing Racial Disparities in the Short Run

Finally, we examine evidence for differential effects by race, which is closely linked with poverty in Brazil. Figures 3.6a and 3.6b suggest that both white and non-white children were more likely to attend school after spraying began, but that the increase for non-white children was larger and less noisy. In particular, the implied static magnitude for the former group was 1.1 p.p. (1.5 percent) whereas for the latter it was 1.9 p.p. (2.6 percent) and more precisely estimated. However, because the 1982 estimate for white children was more negative than for non-white children, the apparent racial difference may simply be an artifact of greater bias toward zero in the white coefficients.

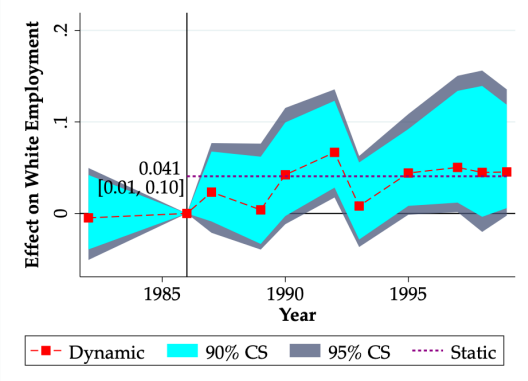
Figures 3.6c and 3.6d also show that employment increased for both groups but somewhat more for non-white adults: the implied white static estimate was 1.0 p.p. (1.5 percent) while the implied non-white one was 1.3 p.p. (1.9 percent). Importantly, the identical differences in outcomes for both racial groups in 1982 and 1986 are consistent with the post-treatment coefficients being accurately estimated. These results thus suggest that Chagas Disease control may have had relatively immediate impacts on the speed of racial convergence in living standards by increasing employment more among non-white Brazilian adults. It also could have set the stage for greater convergence in the future if

Figure 3.6: Differential Short-Run Effects by Race

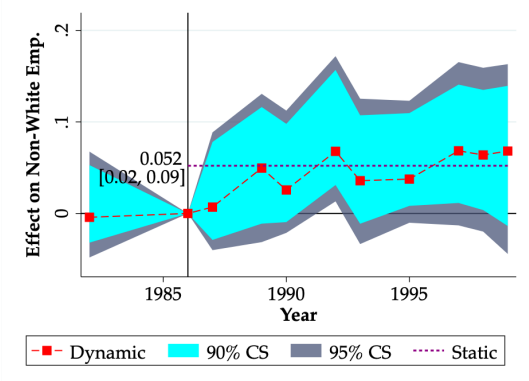


(a) White: Attending School, Ages 8-18

(b) Non-White: Attending School, Ages 8-18



(c) White: Employed, Ages 35-50



(d) Non-White: Employed, Ages 35-50

Notes: Graphs plot dynamic (red squares) and static (dotted purple line) two-way fixed effects estimates for the respective outcomes and racial groups with 90-percent (light blue) and 95-percent (dark blue) wild cluster bootstrap confidence sets for dynamic estimates. Static estimate magnitudes are next to the dotted purple lines with 90-percent wild cluster bootstrap confidence sets below them in brackets. Data are from the 1982-1999 PNAD surveys. Regressions control for fixed effects for state, year, female sex, and racial category (Asian, Black, and Brown) as well as age and age squared. Standard errors are clustered by the 24 consistent 1960-2010 states. Regressions use 389,414 observations for white and 446,017 observations for non-white children, and 648,336 observations for white and 550,144 observations for non-white adults. In the pre-treatment year, 77.7 percent of white and 74.1 percent of non-white children attended school, and 67.2 percent of white and 66.9 percent of non-white adults were employed.

non-white children's human capital increased more as well.

3.6 Long-Run Effects on Labor Market Outcomes

In this section, we examine the long-run effects of triatomine control on adults who were children around the time that spraying began. Our results show that spraying raised incomes and employment rates more among adults who were more exposed to it in childhood. In examining channels, we find small and imprecise increases in years

of schooling, suggesting that this effect does not (primarily) arise through increased educational attainment—a more likely explanation is the direct effect of avoiding the phases of Chagas Disease morbidity. However, these results were markedly different by race: incomes and employment rates rose substantially more for non-white adults despite equivalent (and imprecise) increases in years of schooling. The implication is that vector control increased the speed of racial convergence in Brazil in the long run.

3.6.1 Empirical Strategy

Similar to our approach in the previous section, we compare adults of interest in the 2010 census across municipalities of birth with varying levels of pre-treatment *T. infestans* presence. For dynamic effects, we modify equation (3.5.1) to be

$$y_{i,m,c} = \alpha_m + \gamma_c + \sum_{k \neq 1965} \tau_k \cdot \left(\mathbb{P}[Treat]_m \cdot \mathbb{1}[c = k]\right) + \mathbf{X}_i\beta + \epsilon_{i,m,c}, \qquad (3.6.1)$$
$$\text{for } k \in \{1960, \ldots, 1979\}$$

where the differences are in the subscripts and the omitted period.[12] First, our variables are now at the level of municipality of birth m because of the greater geographic detail in the census data. Nearly two-thirds of the sample live in their municipality of birth, so for these adults $\mathbb{P}[Treat]_m$ is either 0 (55 percent of the sample) or 1 (11 percent). For those living away from their birth municipality, we assign them the probability in the 1980 census that an individual of their sex and race who was living in their state of birth was in a treatment municipality.

Our focus also shifts to birth cohort c. We use 1965 as the omitted cohort, as these adults were 18 years old in 1983. The oldest group we include is the 1960 birth cohort (age 50 in 2010) for 2 reasons. At the time of this census, workers could retire with social security benefits after having paid into the system for 30 to 35 years. Additionally, DDT spraying began in the late 1950s, so we want to avoid comparing cohorts with different levels of exposure to that treatment. The youngest group we include is the 1979 cohort (age 31 in 2010) so that our sample contains prime-age adults with varying exposure to Chagas Disease in childhood and who had already made substantial progress along their lifetime earnings trajectories.

Our assumption is that the 1960 to 1965 cohorts were too old to have experienced the benefits of Chagas Disease control during their childhood years. Given the results in the previous section, it is important to note that these benefits could have come from

[12] We make analogous changes to equation (3.5.2) to estimate the static effect.

both them *as well as their parents* averting the acute phase—the latter may have implied more available resources as a child. On the other hand, these too-old cohorts still may have benefitted from not contracting Chagas Disease in young adulthood, thus avoiding its acute phase in the short run and its chronic determinate phase in the long run. The implication is that null results using this strategy could arise from either the absence of long-run effects from controlling this disease during childhood or the greater importance of its chronic phase for adult outcomes.[13]

The comparisons we make in this section are very similar to those made previously. The τ_k measure the difference in an outcome for a given birth cohort as the probability of having been born in a treatment municipality goes from 0 to 1, relative to the size of that difference for the 1983 cohort. Because the median value is 0 while the 75th percentile remained just under a quarter, we frame our results as an increase in this probability from 0 to 24.7 p.p. and discuss the implied magnitude of the estimate in the text. For inference, we continue to calculate wild cluster bootstrap confidence sets because we cluster standard errors at the state level.

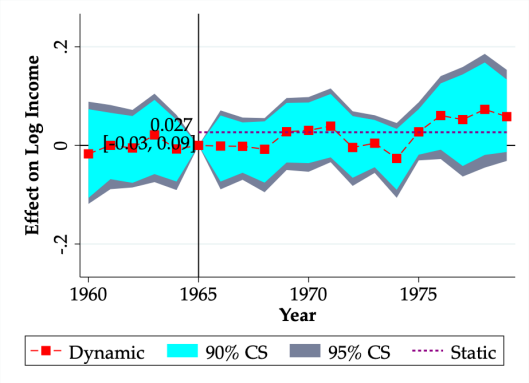
3.6.2 Labor Market Outcomes

Our first outcome of interest is (the natural log of) monthly income for the 1960 to 1979 birth cohorts in 2010. Figure 3.7a shows that while incomes for the adults who turned 18 before the start of spraying evolved in parallel, those who were children during triatomine control experienced slightly greater but imprecise increases the more they were exposed to treatment. The static estimate implied for moving from the median to the 75th-percentile probability of having been born in a treatment municipality is of incomes increasing by 0.7 percent more for treated cohorts.

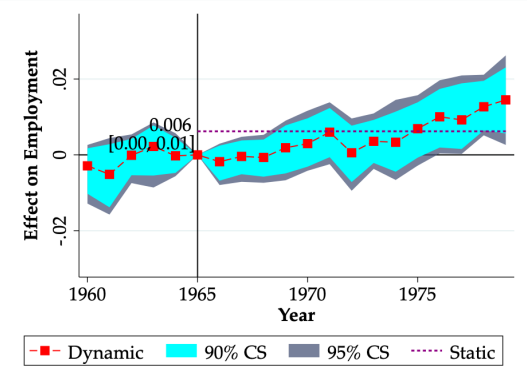
We then turn to studying the employment status of members of these birth cohorts who were in the labor force. After verifying in Appendix Figure C.1 that vector control did not differentially affect labor force participation, Figure 3.7b shows no evidence of differential trends prior to treatment but divergent employment rates after spraying started. The implied static estimate is of an additional 0.1-p.p. increase in employment for treated cohorts. Whether this estimate is of a large relative magnitude depends on how this binary variable is framed—it is 0.2 percent of the employment rate of adults in the labor force from the 1960-65 birth cohorts, which is alternatively 1.7 percent of this group's unemployment rate. Either way, these results provide some evidence that greater exposure to Chagas Disease control in childhood leads to better adult labor market

[13]It also means our long-run results have a somewhat different interpretation from those for childhood exposure to malaria (e.g., Bleakley, 2010b; Lucas, 2010; Cutler et al., 2010), which primarily affects children.

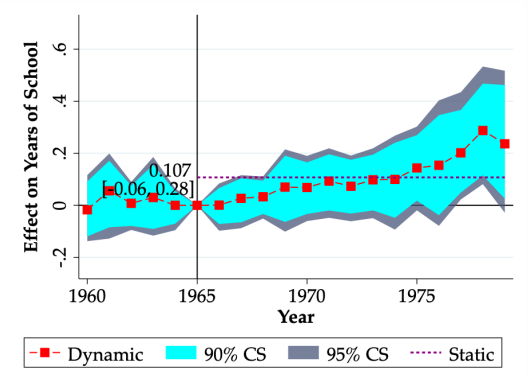
Figure 3.7: Long-Run Effects on Labor Market Outcomes and Educational Attainment



(a) Log Monthly Income



(b) Employed



(c) Years of Schooling

Notes: Graphs plot dynamic (red squares) and static (dotted purple line) two-way fixed effects estimates for the respective outcomes with 90-percent (light blue) and 95-percent (dark blue) wild cluster bootstrap confidence sets for dynamic estimates. Static estimate magnitudes are next to the dotted purple lines with 90-percent wild cluster bootstrap confidence sets below them in brackets. Data are from the IPUMS sample of the 2010 census. Regressions control for state and year fixed effects, female sex, and racial category (Asian, Black, Brown, and Indigenous). Standard errors are clustered by the 24 consistent 1960-2010 states. Regressions use 3,872,397 observations for (a), 4,075,538 for (b), and 5,299,001 for (c). For pre-treatment cohorts, mean log monthly income was 6.242, 91.2 percent of those in the labor force were employed, and mean years of schooling was 6.02.

outcomes.

3.6.3 Assessing Channels: Schooling

Next, we examine the role of schooling in the explanation for our labor market results, especially considering the small and imprecise attendance results in Section 3.5.2. Figure 3.7c shows the absence of pre-treatment differential trends followed by imprecise increases in post-treatment cohorts' completed years. The implied static estimate is of an additional 0.03-year (0.4-percent) increase in years of schooling for cohorts exposed to vector control

as children, though it is noisy.

3.6.4 Reducing Racial Disparities in the Long Run

The imprecision in these results suggests there may be important heterogeneity by race masked by pooling white and non-white Brazilians. Therefore, we study the long-run contributions of Chagas Disease control to reducing racial disparities by examining these outcomes separately for the two groups. Figures 3.8a and 3.8b support our hypothesis of differential effects: while there were small, if any, increases in white incomes for post-treatment cohorts (implied static estimate of 0.2 percent), non-white incomes rose substantially more (implied static estimate of 1.5 percent). In addition, this estimate and the dynamic ones were precise, and the latter imply that non-white cohorts with only a few years of treatment exposure had substantively larger income increases.

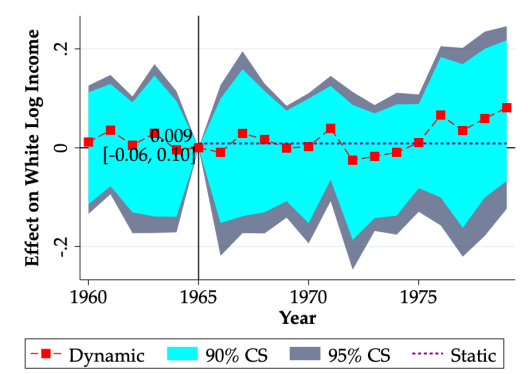
There are also somewhat different effects on employment in adulthood. After verifying that vector control did not have differential outcomes on either racial group's labor force participation in Appendix Figure C.2, we show increases in the implied static estimates of 0.1 p.p. for white rates and for 0.2 p.p for non-white ones in Figures 3.8c and 3.8d. Depending on the framing, these magnitudes are 0.1 percent and 0.3 percent of the respective pre-treatment cohorts' employment rates or 1.5 percent and 2.0 percent of their respective unemployment rates.

The results for years of schooling are more ambiguous. Figures 3.8e and 3.8f show a larger but imprecise increase in absolute terms for white cohorts (implied static estimate of 0.04 years) than non-white ones (0.03 years), but relative to the means for pre-treatment cohorts, the non-white effect is slightly larger (0.5 percent for white cohorts vs 0.6 percent for non-white ones). This absence of substantive differences is consistent with the interpretation of the estimates in Section 3.5.2: namely, after adjusting for the larger downward bias in the white attendance coefficients, there may not have been any differential effects by racial group.

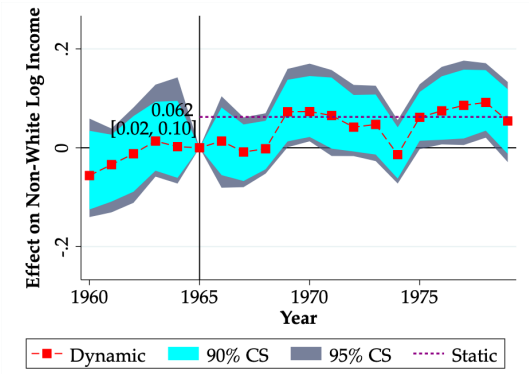
3.6.5 Assessing Channels for Reducing Racial Disparities

Finally, to understand the importance of the schooling channel for the non-white income result, we make a back-of-the-envelope calculation of the upper bound of its contribution. Using the Mincerian return to schooling in Brazil of 15.7 percent reported in Psacharopoulos and Patrinos (2018)—which almost certainly overstates the causal effect of schooling on income—the additional schooling for non-white cohorts induced by vector control can account for at most 0.5 p.p. (one-third) of the 1.5-percent increase in

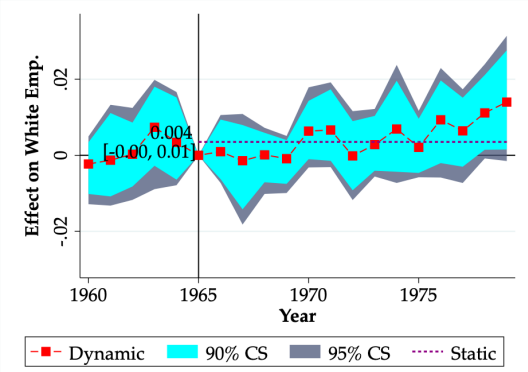
Figure 3.8: Differential Long-Run Effects by Race



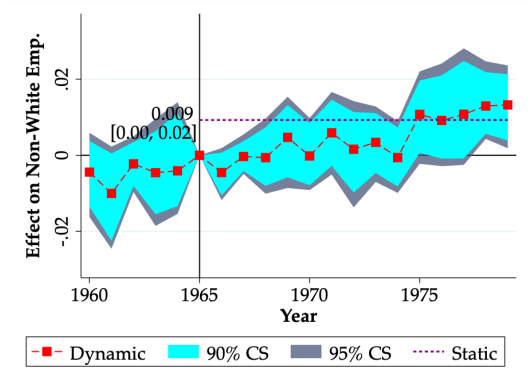
(a) White: Log Monthly Income



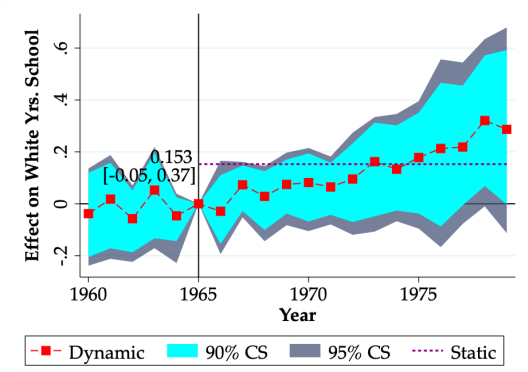
(b) Non-White: Log Monthly Income



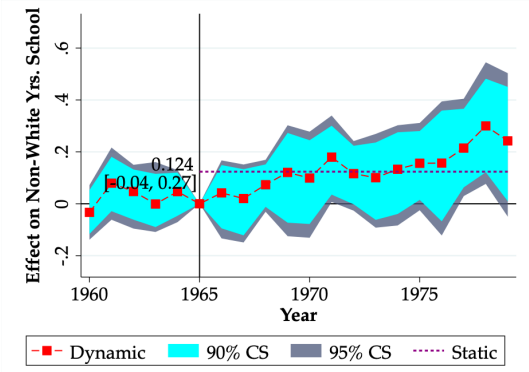
(c) White: Employed



(d) Non-White: Employed



(e) White: Years of Schooling



(f) Non-White: Years of Schooling

Notes: Graphs plot dynamic (red squares) and static (dotted purple line) two-way fixed effects estimates for the respective outcomes and racial groups with 90-percent (light blue) and 95-percent (dark blue) wild cluster bootstrap confidence sets for dynamic estimates. Static estimate magnitudes are next to the dotted purple lines with 90-percent wild cluster bootstrap confidence sets below them in brackets. Data are from the IPUMS sample of the 2010 census. Regressions control for fixed effects for state, year, female sex, and racial category (Asian, Black, and Brown) as well as age and age squared. Standard errors are clustered by the 24 consistent 1960-2010 states. Regressions use the following number of observations: (a) 1,964,673, (b) 1,907,724, (c) 2,044,592, (d) 2,030,946, (e) 2,568,352, (f) 2,730,649. For pre-treatment cohorts, means of the respective outcomes were: (a) 6.58, (b) 5.86, (c) 93.6 percent, (d) 88.5 percent, (e) 6.91 years, (f) 5.10 years.

their incomes. It suggests that avoiding the cardiovascular problems arising from chronic Chagas Disease in adulthood played an important role in explaining differentially larger income increases for non-white Brazilians.

3.7 Long-Run Effects on Public Health Care and Spending

If the morbidity from chronic Chagas Disease was severe enough to keep adults from being productive or even working, reducing it through triatomine control may have impacted more than just individuals' labor market outcomes. As Brazil has the world's largest government-run health care system (the *Sistema Único de Saúde*, or SUS) consuming about 4 percent of its GDP, improvements in adults' cardiovascular health could have had important effects on its public finances. According to SUS data, circulatory system diseases caused one-tenth of the hospitalizations that it paid for since 2010 (over 850,000 per year), which accounted for one-fifth of its spending on hospital care in this period (averaging nearly 1.5 billion 2019 Brazilian *reais* annually, or around 0.1 percent of GDP).

Therefore, in this section we examine the long-run effects of triatomine control on circulatory system-related hospitalizations covered by SUS and the resulting spending. Using a triple-differences strategy comparing circulatory and non-circulatory system-related causes, we show that hospitalizations and spending resulting from the former decreased more in states more exposed to treatment. The implication is that controlling Chagas Disease transmission has yielded substantial benefits for the public health care system and public finances in Brazil.

3.7.1 Data and Empirical Strategy

Our outcomes of interest are (the natural logs of) hospitalizations, person-days spent in the hospital, and spending on hospital care in each state by cause from 1984 to 2019. These data are from the SUS's Hospital Information System (SIH/SUS), and we deflate the last of them so that figures are in constant (log) 2019 Brazilian *reais* (BRL). Given that chronic Chagas Disease manifests primarily as cardiovascular problems, we focus on all diseases of the circulatory system.[14] We combine these data with each state's 1980 share of its population living in treatment municipalities to create a state-level treatment measure as in Section 3.5.

However, there is an additional dimension to our empirical approach that was not necessary in previous sections. The SUS is a heavily decentralized system with transfers of

[14]For 1984 to 1997, we use ICD-9 codes 390-459, and for 1998 onwards, we use ICD-10 codes I00-I99.

responsibilities and funds to state and municipalities (Castro et al., 2018). Therefore, there likely are confounders varying across both state and year (e.g., public health priorities, non-hospital care spending) in violation of the common trends assumption.

To address this complication, we use a triple-differences strategy with all other disease categories as the additional control group. Our assumption is that they are subject to the same state-specific, time-varying factors as circulatory diseases. If it is the case, the triple-difference approach is a valid strategy when the one in previous sections rejects the absence of differential pre-treatment trends for each disease category (Olden and Møen, 2022). The specification we use to estimate dynamic effects is

$$y_{s,t,d} = \alpha_{s,t} + \gamma_{t,d} + \delta_{s,d} + \sum_{k \neq 1999} \tau_k \cdot \left(\mathbb{P}[Treat]_s \cdot \mathbb{1}[t = k] \cdot \mathbb{1}[d = circ] \right) + \epsilon_{s,t,d}, \qquad (3.7.1)$$

where $y_{s,t,d}$ is state s's log outcome in year t due to disease category d, $\alpha_{s,t}$ are state-year fixed effects, $\gamma_{t,d}$ are year-disease category fixed effects, $\delta_{s,d}$ are state-disease category fixed effects, and $\mathbb{1}[d = circ]$ indicates whether the category is circulatory diseases.

This strategy compares first compares the differences in log outcomes due to circulatory diseases in a given year as the probability of a state's population living in a treatment municipality goes from 0 to 1, relative to the size of that difference in 1999 (see below). Then it compares this double-difference estimate to the analogous one for non-circulatory diseases. Because the 25th percentile of this probability was 0 and the 75th percentile was 26.3 p.p., we frame our results as moving between these two values. For inference, we compute wild cluster bootstrap confidence sets after clustering standard errors by the "small" number of states.

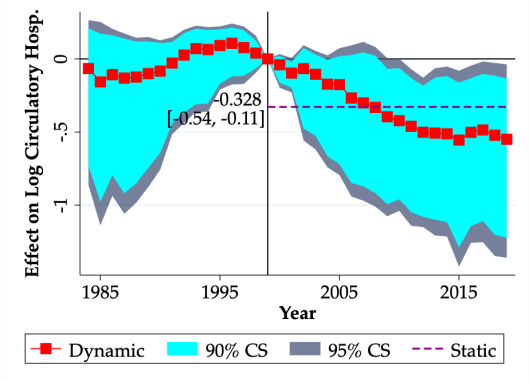
We set 1999 as the reference year because it is 10 years after all treatment municipalities became *T. infestans*-free, and it takes at least 10 years for chronic Chagas Disease to manifest (see Section 3.2). Our hypothesis is that differences within states between circulatory and non-circulatory outcomes should have begun to emerge at or around that point, yielding larger decreases in the former. We also measure this decrease in a static context using a version of equation (3.7.1) analogous to equation (3.5.2).

3.7.2 Circulatory Disease Hospitalizations and Hospital Stay Length

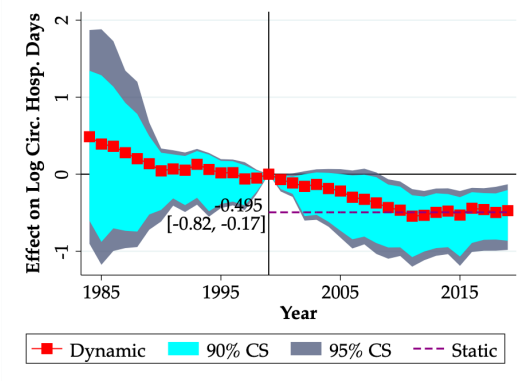
Consistent with our hypothesis, Figure 3.9a shows that the double-difference estimates for log hospitalizations evolved in parallel across disease categories prior to 1999 and then subsequently diverged.[15] In the late 2000s and beyond, the dynamic coefficients implied

[15]In results not shown, we verify that the two categories' double-difference estimates for this and the other two outcomes did not have parallel pre-treatment trends prior to 1999. The implication is that there

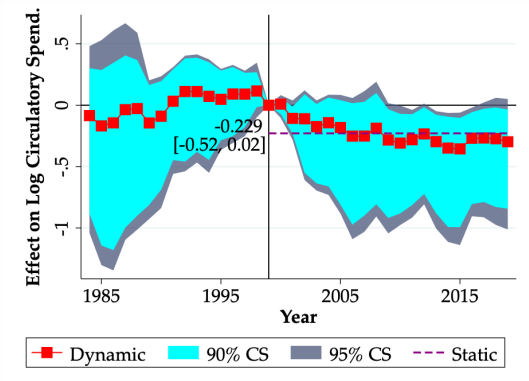
Figure 3.9: Long-Run Effects on Circulatory Disease Hospital Care



(a) Log Circulatory Disease Hospitalizations



(b) Log Circulatory Disease Hospital Days



(c) Log Circulatory Disease Spending

Notes: Graphs plot dynamic (red squares) and static (dotted purple line) triple-difference estimates for the respective outcomes with 90-percent (light blue) and 95-percent (dark blue) wild cluster bootstrap confidence sets for dynamic estimates. Static estimate magnitudes are next to the dotted purple lines with 90-percent wild cluster bootstrap confidence sets below them in parentheses. Data are from DATASUS. Regressions control for state-year, year-disease category, and state-disease category fixed effects. Standard errors are clustered by the 24 consistent 1960-2010 states. Regressions use 1,512 state-year-disease category observations. For pre-treatment years, mean log hospitalizations was 9.94 for circulatory diseases and 12.4 for non-circulatory diseases, mean log person-days in the hospital was 11.9 for circulatory diseases and 14.2 for non-circulatory diseases, and mean log spending was 12.2 for circulatory diseases and 18.1 for non-circulatory diseases.

additional decreases in hospitalizations of at least 8 percent, each of which was significant at the 5-percent level. The economic significance of the implied static coefficient (8.6 percent) as well as its precision suggests that controlling Chagas Disease transmission had a substantial impact on Brazilian health care.

The patterns in log person-days spent in the hospital are highly similar: after pre-1999 trends evolved for the most part in parallel in Figure 3.9b, time spent in the hospital due to circulatory diseases decreased more. By the late 2000s, these differences stabilized at

were indeed state-specific, time-varying confounders necessitating the triple-differences strategy.

an implied magnitude of around 13 percent and achieved 5-percent significance, as did the static coefficient (implied estimate of 13.0 percent). While the latter may be overstated given the positive but noisy coefficients at the start of the sample, these results nonetheless suggest that Brazilians relying on the SUS for health care spent less time in hospitals due to circulatory problems right when we expect this decline to have occurred.

3.7.3 Circulatory Disease Hospital Spending

Because these declines should have translated into a drop in SUS outlays on hospital care, we expect to find a similar pattern in the spending results. Figure C.3f is consistent with this prediction but there is more noise in these estimates. Following pre-1999 coefficients of small magnitudes, the dynamic estimates become consistently more negative and achieve 5-percent significance in the 2010s, although those for last few years of the sample are less precise. But the implied static estimate of an additional 6.0-percent decrease in hospital spending (just outside of 10-percent significance) suggests that SUS has benefited—and continues to benefit—from substantial savings on circulatory disease hospital care spending resulting from triatomine control that occurred more than 3 decades ago.

3.8 Conclusion

Our understanding of the role of disease in explaining differences in economic development between and within countries has mostly been limited to its effects on childhood human capital (usually measured as schooling), which subsequently affects adult incomes for those treated as children. While such impacts are very important for development in the long run, it takes decades to realize their full returns and they are by no means the only long-run economic gains from disease control programs in developing countries. As a result of discounting these benefits and considering those only in this domain, cost-benefit analyses of these campaigns may fail to justify them to policymakers and development practitioners.

However, this paper has shown there were important short-run benefits to Brazil's campaign to control the main vector of Chagas Disease, which has both acute and chronic phases, and important long-run benefits beyond individuals' labor market returns. We found that shortly after spraying began, employment rates increased for *older adults already in the labor force*, which likely resulted in quickly-improved living standards for them and their families. In the long run, vector control raised adult incomes for non-white Brazilians treated as children, potentially helping to increase the speed of racial convergence in a

country with wide disparities in this dimension. We also found small and imprecisely estimated effects on school attendance and eventual educational attainment for both white and non-white adults, suggesting that reducing the cardiovascular morbidity arising from Chagas Disease's chronic phase may be more important in explaining the income result.

Because circulatory diseases result in a substantial share of hospitalizations (10 percent) and spending (20 percent) covered by Brazil's publicly-run health care system, which consumes around 4 percent of GDP, this paper also showed that these outcomes decreased substantially more for circulatory causes than non-circulatory ones in states more exposed to vector control beginning around the time we expected such a difference to arise. We interpret these results as evidence for Chagas Disease control having a significant impact on Brazil's public finances by improving adult health in the long run, which is another important impact not previously examined in the literature.

Taken together, these results present a more complete picture of the economic consequences of Chagas Disease control for developing countries. Whether they generalize beyond this unique malady—which almost exclusively afflicts the Americas, can affect both children and adults in its acute stage, and can cause long-run cardiovascular problems—is an open question we leave to future research. Nonetheless, we believe that this paper has identified novel areas through which health can impact economic development, helping to strengthen justifications for controlling transmission of this neglected tropical disease affecting an estimated 6 million people throughout the Western Hemisphere.

APPENDICES

APPENDIX A

Appendix to Chapter 1

A.1 Data Sources and Variable Definitions

Geographic Traits

- *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 et al. (2014).

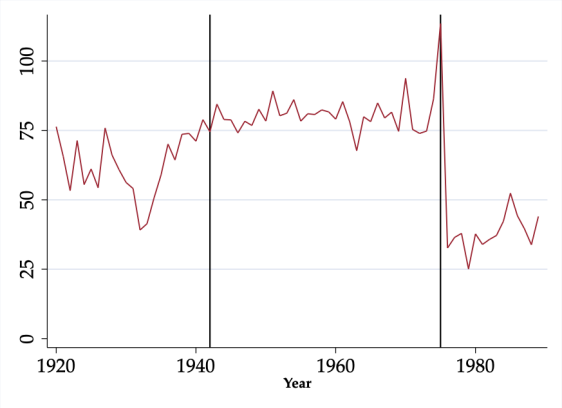
Disease Environment

- *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).

A.2 Additional Figures

A.2.1 Annual Numbers of Witwatersrand Mine Workers from Southern Mozambique

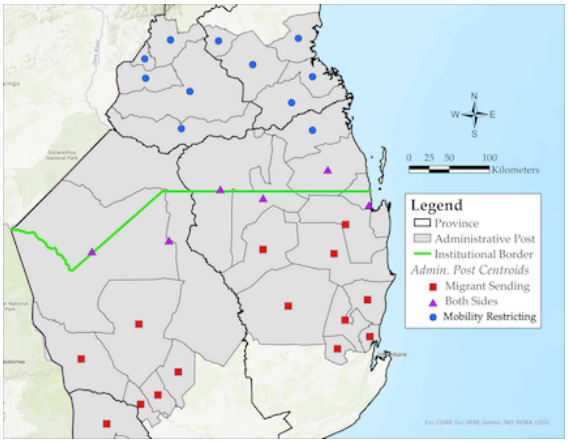
Figure A.1: Southern Mozambican Men Received by Mines (000s), 1920-89



Notes: Data are from the annual reports of the Witwatersrand Native Labour Associated (as cited in Crush et al., 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.

A.2.2 Map of Administrative Posts

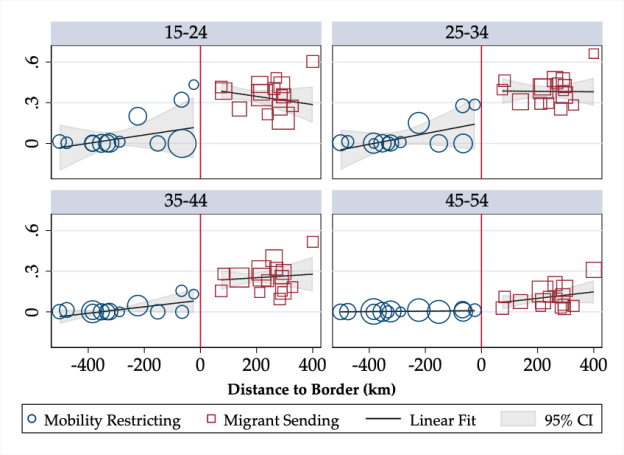
Figure A.2: Map of Administrative Posts



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

A.2.3 Heterogeneity in Circular Migration by Age

Figure A.3: Men's Circular Migration Rates by Age Group



A.3 Additional Tables

A.3.1 HIV Blood Test Refusals

Table A.1: HIV Blood Test Refusals

	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
Mobility Restricting SD	0.094	0.066	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 et al., 2014).

A.3.2 Ruling Out Other HIV Risk Factors

Table A.2: Ruling Out Other HIV Risk Factors

	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 et al., 2014).

A.4 Excluding DHS Clusters within 2 km of the Border

As discussed in Section 1.6.1, 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.

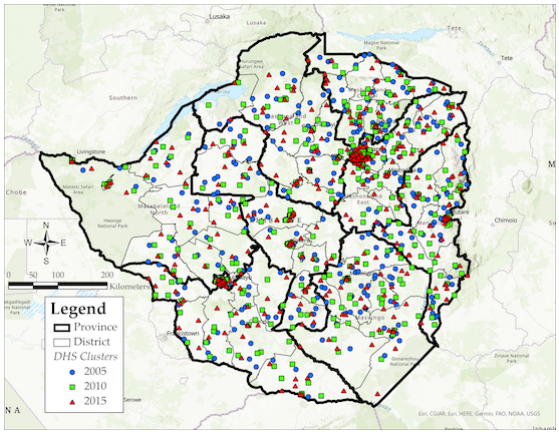
APPENDIX B

Appendix to Chapter 2

B.1 Additional Figures

B.1.1 DHS Clusters

Figure B.1: DHS Clusters

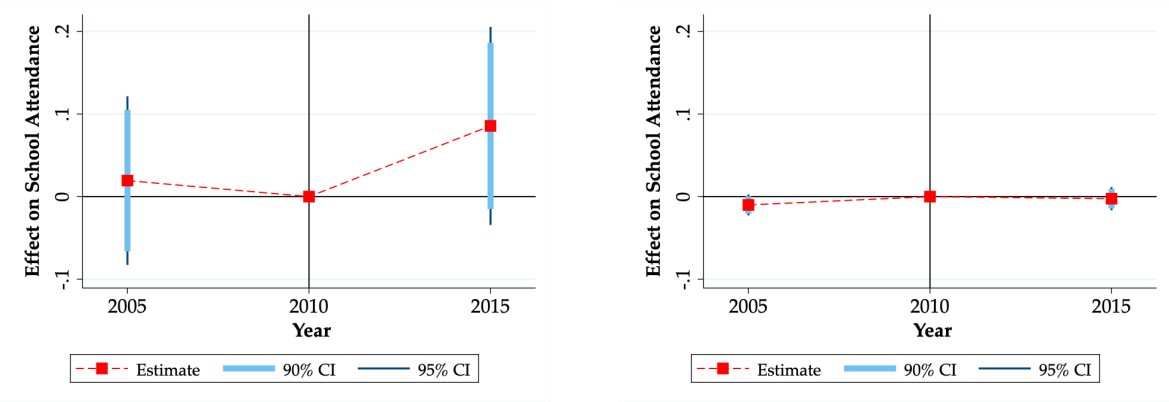


Georeferenced DHS Survey Clusters

Notes: Map shows the locations of survey clusters in the 2005 (blue circles), 2010 (green squared), and 2015 (red triangles) waves of the DHS.

B.1.2 School Attendance

Figure B.2: Secondary School Attendance for Women Ages 18-19

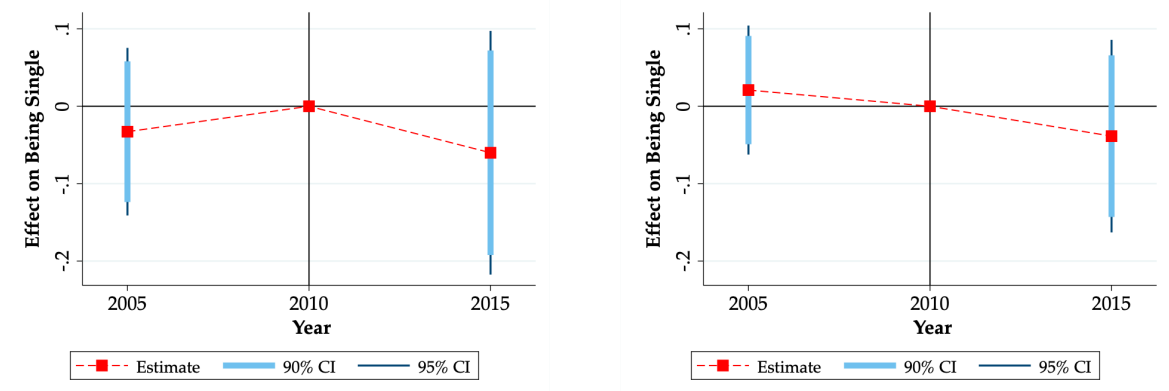


(a) Secondary School: Women Ages 18-19 (b) Higher Education: Women Ages 18-19

Notes: Plots show two-way fixed effect estimates for the respective outcomes. Regressions control for year and district fixed effects, age, age squared, and a quadratic polynomial in latitude and longitude. Standard errors are clustered by the 71 districts in the sample. Regressions use 1,669 observations.

B.1.3 Marriage

Figure B.3: Marriage Rates for Young Women



(a) Single: Women Ages 18-20 (b) Single: Women Ages 21-23

Notes: Plots show two-way fixed effect estimates for being single (i.e., never married or in a union). Regressions control for year and district fixed effects, age, age squared, and a quadratic polynomial in latitude and longitude. Standard errors are clustered by the 71 districts in the sample. Regressions use 2,515 observations for ages 18 to 20 and 2,313 for ages 21 to 23.

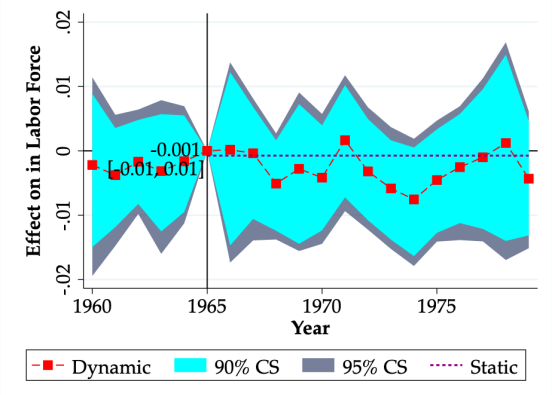
APPENDIX C

Appendix to Chapter 3

C.1 Additional Figures

C.1.1 Labor Force Participation

Figure C.1: Labor Force Participation Rates

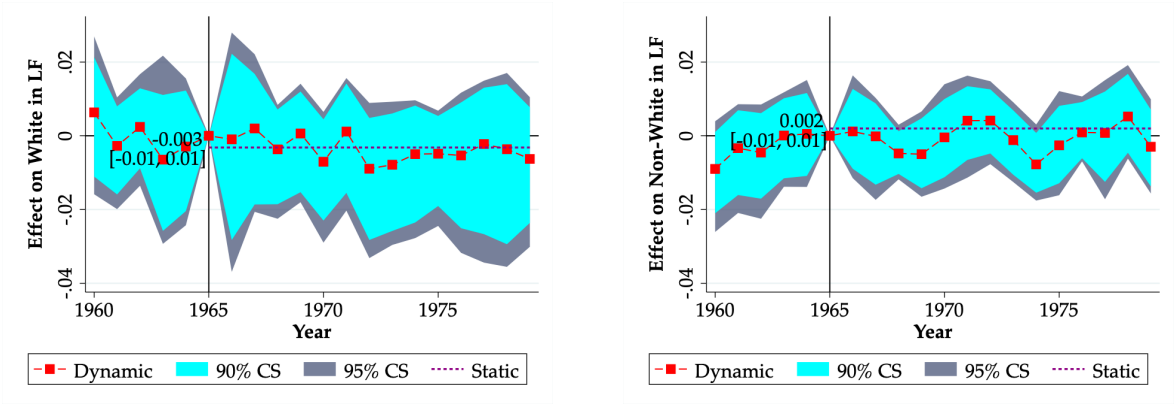


In Labor Force

Notes: Graphs plot dynamic (red squares) and static (dotted purple line) two-way fixed effects estimates for labor force participation with 90-percent (light blue) and 95-percent (dark blue) wild cluster bootstrap confidence sets for dynamic estimates. The static estimate magnitude is next to the dotted purple line with 90-percent wild cluster bootstrap confidence sets below it in brackets. Data are from the IPUMS sample of the 2010 census. Regression controls for state and year fixed effects, female sex, and racial category (Asian, Black, Brown, and Indigenous). Standard errors are clustered by the 24 consistent 1960-2010 states. Regression uses 4,455,390 observations. For pre-treatment cohorts, the labor force participation rate was 73.3 percent.

C.1.2 Labor Force Participation by Race

Figure C.2: Labor Force Participation Rates by Race



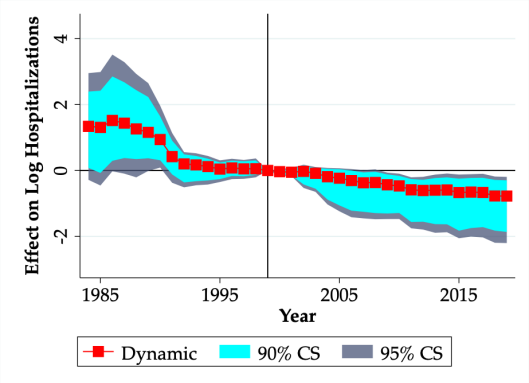
(a) White: In Labor Force

Non-White: In Labor Force

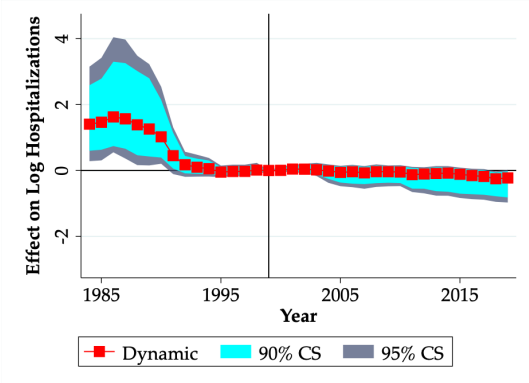
Notes: Graphs plot dynamic (red squares) and static (dotted purple line) two-way fixed effects estimates for labor force participation by race with 90-percent (light blue) and 95-percent (dark blue) wild cluster bootstrap confidence sets for dynamic estimates. Static estimate magnitudes are next to the dotted purple line with 90-percent wild cluster bootstrap confidence sets below them in brackets. Data are from the IPUMS sample of the 2010 census. Regressions control for state and year fixed effects, female sex, and racial category (Asian, Black, Brown, and Indigenous). Standard errors are clustered by the 24 consistent 1960-2010 states. Regressions use 1,983,837 observations for (a) and 2,471,553 observations in (b). For pre-treatment cohorts, labor force participation rates were 76.3 for white adults and 70.6 for non-white adults.

C.1.3 Diff-in-Diff by Disease Category: Hospital Care Outcomes

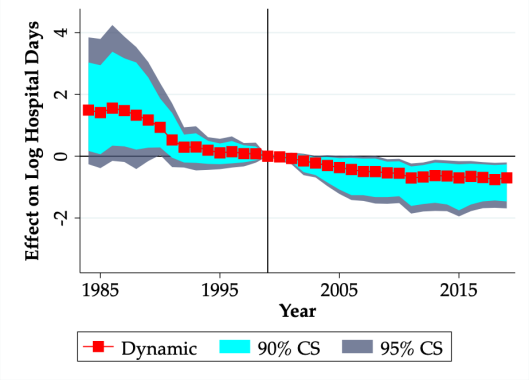
Figure C.3: Diff-in-Diff Estimates for Hospital Care Outcomes by Disease Category



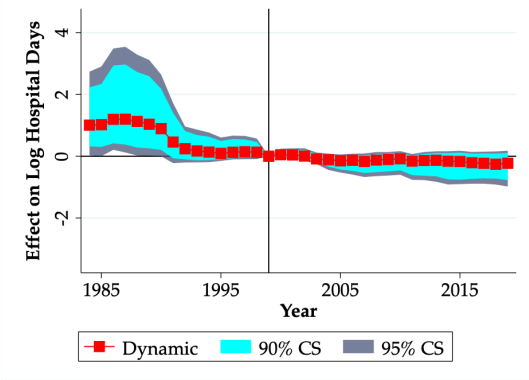
(a) Circulatory Disease: Log Hospitalizations



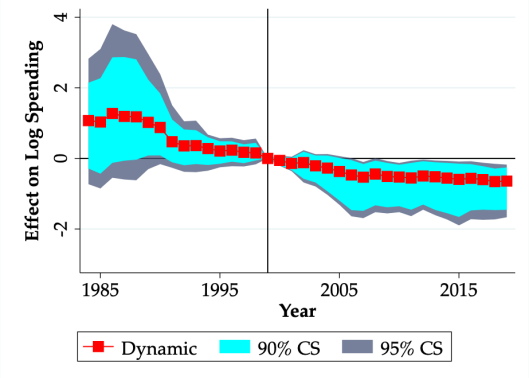
(b) Non-Circ. Disease: Log Hospitalizations



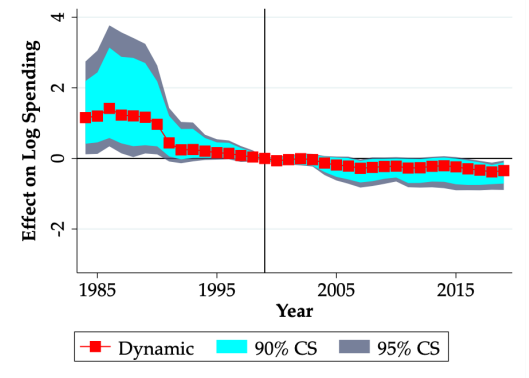
(c) Circulatory Disease: Log Hospital Days



(d) Non-Circ. Disease: Log Hospital Days



(e) Circulatory Disease: Log Spending



(f) Non-Circ. Disease: Log Spending

Notes: Graphs plot dynamic (red squares) and static (dotted purple line) difference-in-difference estimates for the respective outcomes with 90-percent (light blue) and 95-percent (dark blue) wild cluster bootstrap confidence sets. Data are from DATASUS. Regressions control for state and year fixed effects. Standard errors are clustered by the 24 consistent 1960-2010 states. Regressions use 756 state-year observations. See the notes for Figure 3.9 for pre-treatment means for the respective outcomes and disease categories.

BIBLIOGRAPHY

BIBLIOGRAPHY

Aaronson, D., F. Lange, and B. Mazumder (2014): "Fertility Transitions along the Extensive and Intensive Margins," *American Economic Review*, 104, 3701–3724.

Abramitzky, R., L. Boustan, and K. Eriksson (2019): "To the New World and Back Again: Return Migrants in the Age of Mass Migration," *Industrial and Labor Relations Review*, 72, 300–322.

Acemoglu, D. and S. Johnson (2007): "Disease and Development: The Effect of Life Expectancy on Economic Growth," *Journal of Political Economy*, 115, 925–985.

Acemoglu, D., S. Johnson, and J. A. Robinson (2001): "The colonial origins of comparative development: An empirical investigation," *American Economic Review*, 91, 1369–1401.

Acemoglu, D. and J. A. Robinson (2012): *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*, New York: Crown Business.

Agüero, J. M. and P. Bharadwaj (2014): "Do the More Educated Know More about Health? Evidence from Schooling and HIV Knowledge in Zimbabwe," *Economic Development and Cultural Change*, 62, 489–517.

Aiken, A. M., C. Davey, J. R. Hargreaves, and R. J. Hayes (2015): "Re-Analysis of Health and Educational Impacts of a School-Based Deworming Programme in Western Kenya: A Pure Replication," *International Journal of Epidemiology*, 44, 1572–1580.

Alexopoulou, K. and D. Juif (2017): "Colonial State Formation without Integration: Tax Capacity and Labour Regimes in Portuguese Mozambique (1890s-1970s)," *International Review of Social History*, 62, 215–252.

Allina, E. (2012): *Slavery by Any Other Name: African Life under Company Rule in Colonial Mozambique*, Charlottesville, VA: University of Virginia Press.

Almond, D. and J. Currie (2011): "Killing Me Softly: The Fetal Origins Hypothesis," *Journal of Economic Perspectives*, 25, 153–172.

Alsan, M. (2015): "The effect of the TseTse fly on African development," *American Economic Review*, 105, 382–410.

Alsan, M. and M. Wanamaker (2018): "Tuskegee and the health of black men," *Quarterly Journal of Economics*, 133, 407–455.

Amin, S. (1972): "Underdevelopment and Dependence in Black Africa: Origins and Contemporary Forms," *Journal of Modern African Studies*, 10, 503–524.

Anderson, S. (2018): "Legal origins and female HIV," *American Economic Review*, 108, 1407–1439.

Angelucci, M. and D. Bennett (2021): "Adverse Selection in the Marriage Market: HIV Testing and Marriage in Rural Malawi," *Review of Economic Studies*, 88, 2119–2148.

Angrist, J. D. and J.-S. Pischke (2009): *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton: Princeton University Press.

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

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

Ashraf, Q. and O. Galor (2011): "Dynamics and Stagnation in the Malthusian Epoch," *American Economic Review*, 101, 2003–2041.

Ashraf, Q. H., A. Lester, and D. N. Weil (2009): "When Does Improving Health Raise GDP?" *NBER Macroeconomics Annual*, 23, 157–204.

Audet, C. M., J. Burlison, T. D. Moon, M. Sidat, A. E. Vergara, and S. H. Vermund (2010): "Sociocultural and epidemiological aspects of HIV/AIDS in Mozambique," *BMC International Health and Human Rights*, 10, 15.

Baird, S., J. Hamory Hicks, M. Kremer, and E. Miguel (2016): "Worms at Work: Long-Run Impacts of a Child Health Investment," *Quarterly Journal of Economics*, 131, 1637–1680.

Baird, S. J., R. S. Garfein, C. T. McIntosh, and B. Özler (2012): "Effect of a Cash Transfer Programme for Schooling on Prevalence of HIV and Herpes Simplex Type 2 in Malawi: A Cluster Randomised Trial," *Lancet*, 379, 1320–1329.

Banerjee, A. and L. Iyer (2005): "History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India," *American Economic Review*, 95, 1190–1213.

Barakat, S., M. Chard, T. Jacoby, and W. Lume (2002): "The Composite Approach: Research Design in the Context of War and Armed Conflict," *Third World Quarterly*, 23, 991–1003.

Baranov, V. and H.-P. Kohler (2018): "The Impact of AIDS Treatment on Savings and Human Capital Investment in Malawi," *American Economic Journal: Applied Economics*, 10, 266–306.

Becker, G. S. (1991): *A Treatise on the Family*, Cambridge: Harvard University Press.

Behrman, J. A. (2015): "The Effect of Increased Primary Schooling on Adult Women's HIV Status in Malawi and Uganda: Universal Primary Education as a Natural Experiment," *Social Science & Medicine*, 127, 108–115.

Benade, M., B. Nichols, G. Fatti, S. Kuchukhidze, K. Takarinda, N. Mabhena-Ngorima, A. Grimwood, and S. Rosen (2021): "Economic Evaluation of a Cluster Randomized, Non-Inferiority Trial of Differentiated Service Delivery Models of HIV Treatment in Zimbabwe," .

Bertocchi, G. and A. 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.

Björkman Nyqvist, M., L. Corno, D. de Walque, and J. Svensson (2018): "Incentivizing Safer Sexual Behavior: Evidence from a Lottery Experiment on HIV Prevention," *American Economic Journal: Applied Economics*, 10, 287–314.

Black, D. A., S. G. Sanders, E. J. Taylor, and L. J. Taylor (2015): "The Impact of the Great Migration on Mortality of African Americans: Evidence from the Deep South," *American Economic Review*, 105, 477–503.

Bleakley, H. (2007): "Disease and Development: Evidence from Hookworm Eradication in the American South," *Quarterly Journal of Economics*, 122, 73–117.

——— (2010a): "Health, Human Capital, and Development," *Annual Review of Economics*, 2, 283–310.

——— (2010b): "Malaria Eradication in the Americas: A Retrospective Analysis of Childhood Exposure," *American Economic Journal: Applied Economics*, 2, 1–45.

Bleakley, H. and F. Lange (2009): "Chronic Disease Burden and the Interaction of Education, Fertility, and Growth," *Review of Economics and Statistics*, 91, 52–65.

Bocchi, E. A., G. Guimarães, F. Tarasoutshi, G. Spina, S. Mangini, and F. Bacal (2009): "Cardiomyopathy, Adult Valve Disease and Heart Failure in South America," *Heart*, 95, 181–189.

Bor, J., F. Tanser, M.-L. Newell, and T. Bärnighausen (2012): "In a Study of a Population Cohort in South Africa, HIV Patients on Antiretrovirals Had Nearly Full Recovery of Employment," *Health Affairs*, 31, 1459–1469.

Cagé, J. and V. Rueda (2020): "Sex and the Mission: The Conflicting Effects of Early Christian Missions on HIV in sub-Saharan Africa," *Journal of Demographic Economics*, 86, 213–257.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014): "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 82, 2295–2326.

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

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008): "Bootstrap-based improvements for inference with clustered errors," *Review of Economics and Statistics*, 90, 414–427.

Case, A., A. Fertig, and C. Paxson (2005): "The Lasting Impact of Childhood Health and Circumstance," *Journal of Health Economics*, 24, 365–389.

Case, A. and C. Paxson (2013): "HIV Risk and Adolescent Behaviors in Africa," *American Economic Review*, 103, 433–438.

Castro, M. C., A. Massuda, G. Almeida, N. A. Menezes-Filho, M. V. Andrade, K. V. M. d. S. Noronha, R. Rocha, J. Macinko, T. Hone, R. Tasca, L. Giovanella, A. M. Malik, H. Werneck, L. A. Fachini, and R. Atun (2018): "Brazil's Unified Health System: The First 30 Years and Prospects for the Future," *Lancet*, 394, 345–356.

Cattaneo, M. D., N. Idrobo, and R. Titiunik (2019): *A Practical Introduction to Regression Discontinuity Designs: Foundations*, Cambridge, UK: Cambridge University Press.

Chen, C. Y., R. C. Ballard, C. Beck-Sague, Y. Dangor, F. Radebe, S. Schmid, J. Weiss, V. Tshabalala, G. Fehler, Y. Htun, and S. Morse (2000): "Human Immunodeficiency Virus Infection and Genital Ulcer Disease in South Africa," *Sexually Transmitted Diseases*, 27, 21–29.

Chiappori, P.-A., M. Costa Dias, and C. Meghir (2018): "The Marriage Market, Labor Supply, and Education Choice," *Journal of Political Economy*, 126, S26–S72.

Chiappori, P.-A., M. Iyigun, and Y. Weiss (2009): "Investment in Schooling and the Marriage Market," *American Economic Review*, 99, 1689–1713.

Clarence-Smith, G. (1985): *The third Portuguese empire, 1825-1975: A study in economic imperialism*, Manchester: Manchester University Press.

Colella, F., R. Lalive, S. O. Sakalli, and M. Thoenig (2020): "Inference with Arbitrary Clustering," .

Colley, D. G., A. L. Bustinduy, W. E. Secor, and C. H. King (2014): "Human Schistosomiasis," *Lancet*, 383, 2253–2264.

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

Corno, L., N. Hildebrandt, and A. Voena (2020): "Age of Marriage, Weather Shocks, and the Direction of Marriage Payments," *Econometrica*, 88, 879–915.

Coura, J. R. and J. C. P. Dias (2009): "Epidemiology, Control and Surveillance of Chagas Disease - 100 Years after its Discovery," *Memórias do Instituto Oswaldo Cruz*, 104, 31–40.

Croke, K., J. Hamory Hicks, E. Hsu, M. Kremer, R. Maertens, E. Miguel, and W. Więcek (2022): "Meta-Analysis and Public Policy: Reconciling the Evidence on Deworming," .

Crush, J., A. Jeeves, and D. Yudelman (1991): *South Africa's labor empire*, Boulder, CO: Westview Press.

Currie, J. (2009): "Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development," *Journal of Economic Literature*, 47, 87–122.

Currie, J. and D. Almond (2011): "Human Capital Development before Age 5," in *Handbook of Labor Economics*, ed. by D. Card and O. Ashenfelter, Amsterdam: North-Holland, vol. 4B, 1315–1486.

Cutler, D., W. Fung, M. Kremer, M. Singhal, and T. Vogl (2010): "Early-Life Malaria Exposure and Adult Outcomes: Evidence from Malaria Eradication in India," *American Economic Journal: Applied Economics*, 2, 72–94.

Danielson, J. J. and D. B. Gesch (2011): "Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010)," Tech. Rep. U.S. Geological Survey Open-File Report 2011–1073, Reston, VA.

Darch, C. (1983): "Notas sobre Fontes Estatísticas Oficiais Referente à Economia Colonial Moçambicana: Uma Crítica Geral," *Estudos Moçambicanos*, 4, 103–125.

Davey, C., A. M. Aiken, R. J. Hayes, and J. R. Hargreaves (2015): "Re-Analysis of Health and Educational Impacts of a School-Based Deworming Programme in Western Kenya: A Statistical Replication of a Cluster Quasi-Randomized Stepped-Wedge Trial," *International Journal of Epidemiology*, 44, 1581–1592.

De Neve, J.-W., G. Fink, S. V. Subramanian, S. Moyo, and J. Bor (2015): "Length of Secondary Schooling and Risk of HIV Infection in Botswana: Evidence from a Natural Experiment," *Lancet Global Health*, 3, e470–e477.

de Oliveira, T., A. B. M. Kharsany, T. Gräf, C. Cawood, D. Khanyile, A. Grobler, A. Puren, S. Madurai, C. Baxter, Q. Abdool Karim, and S. 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, E41–E50.

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

Dell, M. and B. A. Olken (2020): "The Development Effects of the Extractive Colonial Economy: The Dutch Cultivation System in Java," *Review of Economic Studies*, 87, 164–203.

Derenoncourt, E. (2021): "Can You Move to Opportunity? Evidence from the Great Migration," .

Dias, J. C. P. (1987): "Control of Chagas Disease in Brazil," *Parasitology Today*, 3, 336–341.

Dinkelman, T. and M. Mariotti (2016): "The long-run effects of labor migration on human capital formation in communities of origin," *American Economic Journal: Applied Economics*, 8, 1–35.

Direcção Provincial dos Serviços de Estatística (1966): *III Recenseamento Geral da População na Província de Moçambique*, Lourenço Marques: Instituto Nacional de Estatística.

Duflo, E., P. Dupas, and M. Kremer (2015): "Education, HIV, and Early Fertility: Experimental Evidence from Kenya," *American Economic Review*, 105, 2757–2797.

Dunkle, K. L., R. K. Jewkes, H. C. Brown, G. E. Gray, J. A. McIntryre, and S. D. Harlowa (2004): "Transactional Sex among Women in Soweto, South Africa: Prevalence, Risk Factors and Association with HIV Infection," *Social Science & Medicine*, 59, 1581–1592.

Dupas, P. (2011): "Do Teenagers Respond to HIV Risk Information? Evidence from a Field Experiment," *American Economic Journal: Applied Economics*, 3, 1–34.

Dwyer-Lindgren, L., M. A. Cork, A. Sligar, K. M. Steuben, K. F. Wilson, N. R. Provost, B. K. Mayala, J. D. VanderHeide, M. L. Collison, J. B. Hall, M. H. Biehl, A. Carter, T. Frank, D. Douwes-Schultz, R. Burstein, D. C. Casey, A. Deshpande, L. Earl, C. El Bcheraoui, ..., and S. I. Hay (2019): "Mapping HIV prevalence in sub-Saharan Africa between 2000 and 2017," *Nature*, 570, 189–193.

Evans, M., K. Risher, N. Zungu, O. Shisana, S. Moyo, D. C. Celentano, B. Maughan-Brown, and T. 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, 21310.

Fernández Cebrián, P. (2021): "Withdrawal of the State: The Provision of Primary Schooling in Mozambique under the Indigenato," .

Friedman, J. F., H. K. Kanzaria, and S. T. McGarvey (2005): "Human Schistosomiasis and Anemia: The Relationship and Potential Mechanisms," *Trends in Parasitology*, 21, 386–392.

Frimmel, H. E. (2019): "The Witwatersrand Basin and Its Gold Deposits," in *The Archean Geology of the Kaapvaal Craton, Southern Africa*, ed. by A. Kröner and A. Hoffman, Cham, CHE: Springer, Regional Geology Reviews.

Fuller, C. E. (1955): "An Ethnohistoric Study of Continuity and Change in Gwambe Culture," PhD dissertation, Northwestern University, Evanston, IL.

Galor, O. and D. N. Weil (2000): "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond," *American Economic Review*, 90, 806–828.

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

Gengenbach, H. (2010): *Binding memories: Women and makers and tellers of history in Magude, Mozambique*, New York: Columbia University Press.

Gertler, P., J. Heckman, R. Pinto, A. Zanolini, C. Vermeersch, S. Walker, S. M. Chang, and S. Grantham-McGregor (2014): "Labor Market Returns to an Early Childhood Stimulation Intervention in Jamaica," *Science*, 344, 998–1001.

GiveWell (2017): "Deworming Might Have Huge Impact, but Might Have Close to Zero Impact," .

Glewwe, P. and E. A. Miguel (2007): "The Impact of Child Health and Nutrition on Education in Less Developed Countries," in *Handbook of Development Economics*, Amsterdam: North-Holland, vol. 4, 3561–3606.

Goodman-Bacon, A. (2021): "Differences-in-Differences with Variation in Treatment Timing," *Journal of Econometrics*, 225, 254–277.

Great Britain Foreign Office (1901): "Portuguese degree, extending the concessions granted to the Mozambique Company - Lisbon, December 22, 1893," in *British and foreign state papers, 1896-1897*, ed. by A. H. Oakes and W. Maycock, London: His Majesty's Stationery Office, vol. 89, 601–602.

Greenwood, J., N. Guner, and G. Vandenbroucke (2017): "Family Economics Writ Large," *Journal of Economic Literature*, 55, 1346–1434.

Greenwood, J., P. Kircher, C. Santos, and M. Tertile (2019): "An Equilibrium Model of the African HIV/AIDS Epidemic," *Econometrica*, 87, 1081–1113.

Griffith, D. A. and G. Arbia (2010): "Detecting Negative Spatial Autocorrelation in Georeferenced Random Variables," *International Journal of Geographical Information Science*, 24, 417–437.

Guthrie, Z. K. (2016): "Forced Volunteers: The Complexities of Coercion in Central Mozambique, 1941-1961," *International Journal of African Historical Studies*, 49, 195–212.

——— (2018): *Bound for Work: Labor, Mobility, and Colonial Rule in Central Mozambique, 1940-1965*, Charlottesville, VA: University of Virginia Press.

Habyarimana, J., B. Mbakile, and C. Pop-Eleches (2010): "The Impact of HIV/AIDS and ARV Treatment on Worker Absenteeism: Implications for African Firms," *Journal of Human Resources*, 45, 809–839.

Hamory, J., E. Miguel, M. Walker, M. Kremer, and S. Baird (2021): "Twenty-Year Economic Impacts of Deworming," *Proceedings of the National Academy of Sciences*, 118, e2023185118.

Hamory Hicks, J., M. Kremer, and E. Miguel (2015): "Commentary: Deworming Externalities and Schooling Impacts in Kenya: A Comment on Aiken et al. (2015) and Davey et al. (2015)," *International Journal of Epidemiology*, 44, 1593–1596.

Hargreaves, J. R., A. M. Aiken, C. Davey, and R. J. Hayes (2015): "Authors' Response to: Deworming Externalities and School Impacts in Kenya," *International Journal of Epidemiology*, 44, 1596–1599.

Harries, P. (1983): "Labour Migration from Mozambique to South Africa; with Special Reference to the Delagoa Bay Hinterland, c. 1862 to 1897," PhD dissertation, School of Oriental and African Studies, University of London, London.

——— (1994): *Work, culture, and identity: Migrant laborers in Mozambique and South Africa, c. 1860-1910*, Portsmouth, NH: Heinemann.

Harris, M. (1959): "Labour Emigration among the Moçambique Thonga: Cultural and Political Factors," *Africa: Journal of the International African Institute*, 29, 50–66.

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

Havik, P. J. (2013): "Colonial Administration, Public Accounts and Fiscal Extraction: Policies and Revenues in Portuguese Africa (1900-1960)," *African Economic History*, 41, 159–221.

Helgesson, A. (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.

Hotez, P. J., D. A. P. Bundy, K. Beegl, S. Brooker, L. Drake, N. de Silva, A. Montresor, D. Engels, M. Jukes, L. Chitsulo, J. Chow, R. Laxminarayan, C. M. Michaud, R. Correa-Oliveira, X. Shu-Hua, A. Fenwick, and L. Savioli (2006): "Helminth Infections: Soil-Transmitted Helminth Infections and Schistosomiasis," in *Disease Control Priorities in Developing Countries*, ed. by J. Bethony, D. T. Jamison, J. G. Breman, A. R. Measham, G. Alleyne, M. Claeson, D. B. Evans, P. Jha, A. Mills, and P. Musgrove, Washington, DC: World Bank and Oxford University Press, 467–482, 2 ed.

Hoynes, H., D. W. Schanzenbach, and D. Almond (2016): "Long-Run Impacts of Childhood Access to the Safety Net," *American Economic Review*, 106, 903–934.

Iliffe, J. (2006): *The African AIDS epidemic: A history*, Athens, OH: Ohio University Press.

Instituto Nacional de Estatística and ICF (2019): *Moçambique Inquérito Nacional sobre Indicadores de Malária [data set]*, Rockville, MD.

Isaacman, A., M. Stephen, Y. Adam, M. J. Homen, E. Macamo, and A. 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, 581–615.

Jayachandran, S. and A. Lleras-Muney (2009): "Life Expectancy and Human Capital Investments: Evidence from Maternal Mortality Declines," *Quarterly Journal of Economics*, 124, 349–397.

Junod, H. A. (1912): *The life of a South African tribe*, London: David Nutt.

Kates, J., A. Wexler, E. Lief, and UNAIDS (2020): *Donor Government Funding for HIV in Low- and Middle-Income Countries in 2019*, San Francisco: Kaiser Family Foundation.

Kayuni, S., F. Lampiao, P. Makuala, L. Juziwelo, E. J. Lacourse, J. Reinhard-Rupp, P. D. C. Leutscher, and J. R. Stothard (2019): "A Systematic Review with Epidemiological Update of Male Genital Schistosomiasis (MGS): A Call for Integrated Case Management Across the Health System in Sub-Saharan Africa," *Parasite Epidemiology and Control*, 4, e00077.

Kelly, M. (2020): "Understanding Persistence," .

Khan, M. G. (2011): "Chagas Disease," in *Encyclopedia of Heart Diseases*, New York: Springer, 295–299, 2 ed.

Khanna, G., C. Theoharides, and D. Yang (2020): "Abundance from Abroad: Migrant Income and Long-Run Economic Development," .

Kiszewski, A., A. Mellinger, A. Spielman, P. Malaney, S. E. Sachs, and J. Sachs (2004): "A global index representing the stability of malaria transmission," *American Journal of Tropical Medicine and Hygiene*, 70, 486–498.

Kjetland, E., P. D. Ndhlovu, E. Gomo, T. Mduluza, N. Midzi, L. Gwanzura, P. R. Mason, L. Sandvik, H. Friis, and S. G. Gundersen (2006): "Association between Genital Schistosomiasis and HIV in Rural Zimbabwean Women," *AIDS*, 20, 593–600.

Leclerc-Madlala, S. (2008): "Age-Disparate and Intergenerational Sex in Southern Africa: The Dynamics of Hypervulnerability," *AIDS*, 22, S17–S25.

Leutscher, P. D. C., M. Pedersen, C. Raharisolo, J. S. Jensen, S. Hoffman, I. Lisse, S. R. Ostrowski, C. M. Reimert, P. Mauclere, and H. Ullum (2005): "Increased Prevalence of Leukocytes and Elevated Cytokine Levels in Semen from Schistosoma haematobium-Infected Individuals," *Journal of Infectious Diseases*, 191, 1639–1647.

Lowes, S. and E. Montero (2021a): "Concessions, Violence, and Indirect Rule: Evidence from the Congo Free State," .

——— (2021b): "The Legacy of Colonial Medicine in Central Africa," *American Economic Review*, 111, 1284–1314.

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

Mabaso, M., L. Mlangeni, L. Makola, O. Oladimeji, I. Naidoo, Y. Naidoo, B. Chibi, K. Zuma, and L. 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.

Maccini, S. and D. Yang (2009): "Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall," *American Economic Review*, 99, 1006–1026.

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

Mduluza, T., C. Jones, D. N. M. Osakunor, R. Lim, J. K. Kuebel, I. Phiri, P. Manangazira, P. Tagwireyi, and F. Mutapi (2020): "Six Rounds of Annual Praziquantel Treatment during a National Helminth Control Program Significantly Reduced Schistosome Infection and Morbidity Levels in a Cohort of Schoolchildren in Zimbabwe," *PLOS Neglected Tropical Diseases*, 14, e0008388.

Michalopoulos, S. and E. Papaioannou (2014): "National institutions and subnational development," *Quarterly Journal of Economics*, 129, 151–213.

——— (2016): "The long-run effects of the Scramble for Africa," *American Economic Review*, 106, 1802–1848.

Midzi, N., T. Mduluza, M. J. Chimbari, C. Tshuma, L. Charimari, G. Mhlanga, P. Manangazira, S. M. Munyati, I. Phiri, and S. L. Mutambu, et al (2014): "Distribution of Schistosomiasis and Soil Transmitted Helminthiasis in Zimbabwe: Towards a National Plan of Action for Control and Elimination," *PLOS Neglected Tropical Diseases*, 8, e3014.

Midzi, N., T. Mduluza, B. Mudenge, L. Foldager, and P. D. C. Leutscher (2017): "Decrease in Seminal HIV-1 RNA Load after Praziquantel Treatment of Urogenital Schistosomiasis Coinfection in HIV-Positive Men: An Observational Study," *Open Forum Infectious Diseases*, 4, ofx199.

Miguel, E. and M. Kremer (2004): "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica*, 72, 159–217.

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

Ministério de Saúde, Instituto Nacional de Estatística, and ICF International (2013): *Moçambique Inquérito Demográfico e de Saúde 2011 [data set]*, Calverton, MD.

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 [data set]*, Calverton, MD.

Ministério do Ultramar (1959): "Carta de Moçambique," .

Minnesota Population Center (2020): *Integrated Public Use Microdata Series, International: Version 7.3 [data set]*, Minneapolis, MN.

Moran, P. A. P. (1950): "Notes on Continuous Stochastic Phenomena," *Biometrika*, 37, 17–23.

Morier-Genoud, E. (2019): *Catholicism and the Making of Politics in Central Mozambique, 1940-1986*, Rochester: University of Rochester Press.

Mozambique News Agency (1999): "Huge Growth at Vilankulo Airport," *AIM Reports*.

——— (2000): "Tourism Brings in $32 Million," *AIM Reports*.

——— (2011): "President Inaugurates New Terminal at Vilankulo Airport," *AIM Reports*.

Murdock, G. P. (1959): *Africa: Its Peoples and Their Culture History*, New York: McGraw-Hill Book Company.

Méndez-Chacón, E. and D. Van Patten (2021): "Multinationals, Monopsony, and Local Development: Evidence from the United Fruit Company," .

Ndeffo Mbah, M. L., E. F. Kjetland, K. E. Atkins, E. M. Poolman, E. W. Orenstein, L. A. Meyers, J. P. Townsend, and A. P. Galvani (2013): "Cost-Effectiveness of a Community-Based Intervention for Reducing the Transmission of Schistosoma haematobium and HIV in Africa," *Proceedings of the National Academy of Sciences*, 110, 7952–7957.

Newitt, M. (1995): *A History of Mozambique*, Bloomington, IN: University of Indiana Press.

North, D. (1990): *Institutions, Institutional Change and Economic Performance*, New York: Cambridge University Press.

Nunes, M. C. P., A. Beaton, H. Acquatella, C. Bern, A. F. Bolger, L. E. Echeverría, W. O. Dutra, J. Gascon, C. A. Morillo, J. Oliveira-Filho, A. L. P. Ribeiro, J. A. Marin-Neto, and On behalf of the American Heart Association Rheumatic Fever, Endocarditis and Kawasaki Disease Committee of the Council on Cardiovascular Disease in the Young; Council on Cardiovascular and Stroke Nursing; and Stroke Council (2018): "Chagas Cardiomyopathy: An Update of Current Clinical Knowledge and Management: A Scientific Statement from the American Heart Association," *Circulation*, 138, e169–e209.

Nunn, N. (2020): "The Historical Roots of Economic Development," *Science*, 367, eaaz9986.

Olden, A. and J. Møen (2022): "The Triple Difference Estimator," *Econometrics Journal*, Forthcoming.

Oster, E. (2012): "Routes of Infection: Exports and HIV Incidence in Sub-Saharan Africa," *Journal of the European Economic Association*, 10, 1025–1058.

Oum, S., A. Carbaugh, and J. Kates (2021): "Funding for Key HIV Commodities in PEPFAR Countries," *Kaiser Family Foundation*.

Ozier, O. (2018): "Exploring Externalities to Estimate the Long-Term Effects of Early Childhood Deworming," *American Economic Journal: Applied Economics*, 10, 235–262.

Peters, M. and A. Siow (2002): "Competing Premarital Investments," *Journal of Political Economy*, 110, 592–608.

Portugal (1900): *Collecção Official de Legislação Portugueza: Anno de 1899*, Lisboa: Imprensa Nacional.

Psacharopoulos, G. and H. A. Patrinos (2018): *Returns to Investment in Education: A Decennial Review of the Global Literature*, Policy Research Working Paper No. 8402, Washington, DC: World Bank.

Quinn, T. C., M. J. Wawer, N. K. Sewankambo, D. Serwadda, C. Li, F. Wabwire-Mangen, M. O. Meehan, T. Lutalo, and R. 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.

Rassi, A., J. C. P. Dias, J. A. Marin-Neto, and A. Rassi (2009): "Challenges and Opportunities for Primary, Secondary, and Tertiary Prevention of Chagas' Disease," *Heart*, 95, 524–534.

Rassi, A., A. Rassi, and W. C. Little (2000): "Chagas' Heart Disease," *Clinical Cardiology*, 23, 883–889.

Reniers, G. and R. Tfaily (2008): "Polygyny and HIV in Malawi," *Demographic Research*, 19, 1811–1830.

Repartição Nacional de Estatística (1942): *Censo da População em 1940*, vol. 2-5, Lourenço Marques: Imprensa Nacional de Moçambique.

Reynoso, A. (2021): "Polygamy, Co-Wives' Complementarities, and Intrahousehold Inequality," .

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

Roberts, R. (2017): "Coerced Labor in Twentieth-Century Africa," in *The Cambridge World History of Slavery*, ed. by D. Eltis, S. L. Engerman, S. Drescher, and D. Richardson, Cambridge: Cambridge University Press, 583–609.

Robinson, J. and E. Yeh (2011): "Transactional Sex as a Response to Risk in Western Kenya," *American Economic Journal: Applied Economics*, 3, 35–64.

Rocha, R. and R. R. Soares (2010): "Evaluating the Impact of Community-Based Health Interventions: Evidence from Brazil's Family Health Program," *Health Economics*, 19, 126–158.

Ross, E. A. (1925): *Report on Employment of Native Labor in Portuguese Africa*, New York: The Abbott Press.

Saldanha, F. (1940): *Carta da Colónia de Moçambique: Segundo a Sua Actual Divisão Administrativa [map]*, Lourenço Marques: Direcção dos Serviços de Agrimensura.

Schaefer, R., S. Gregson, J. W. Eaton, O. Mugurungi, R. Rhead, A. Takaruza, R. Maswera, and C. Nyamukapa (2017): "Age-disparate relationships and HIV incidence in adolescent girls and young women: Evidence from Zimbabwe," *AIDS*, 31, 1461–1470.

Schneider, U., A. Becker, P. Finger, E. Rustemeier, and M. 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," .

Schofield, C. J. (1981): "Chagas Disease, Triatomine Bugs, and Blood-Loss," *The Lancet*, 317, 1316.

——— (1988): "Biosystematics of the Triatominae," in *Biosystematics of Haematophagous Insects*, ed. by M. W. Service, Oxford, UK: Clarendon Press.

Schofield, C. J. and J. C. P. Dias (1999): "The Southern Cone Initiative against Chagas Disease," in *Advances in Parasitology*, ed. by J. R. Baker, R. Muller, and D. Rollinson, San Diego: Academic Press, vol. 42, 1–27.

Silveira, A. C. (2011): "O Inquérito Triatomínico (1975-1983)," *Revista da Sociedade Brasileira de Medicina Tropical*, 44, 26–32.

Siow, A. (1998): "Differential Fecundity, Markets, and Gender Roles," *Journal of Political Economy*, 106, 334–354.

Smith, A. K. and G. C. Smith (1985): "Angola and Mozambique, 1870-1905," in *The Cambridge history of Africa*, ed. by R. Oliver and G. N. Sanderson, Cambridge: Cambridge University Press, vol. 6, 493–521.

Stichter, S. (1985): *Migrant Laborers*, Cambridge, UK: Cambridge University Press.

Tangeni, T. (2007): "Contextual Factors Associated with Treatment-Seeking and Higher-Risk Sexual Behavior in Botswana among Men with Symptoms of Sexually Transmitted Infections," *African Journal of AIDS Research*, 6, 261–269.

Tanser, F., T. Bärnighausen, L. Hund, G. P. Garnett, N. McGrath, and M.-L. 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, 247–255.

Taylor-Robinson, D. C., N. Maayan, S. Donegan, M. Chaplin, and P. Gerner (2019): "Public Health Deworming Programmes for Soil-Transmitted Helminths in Children Living in Endemic Areas (Review)," *Cochrane Database of Systematic Reviews*, 9, CD000371.

Tertilt, M. (2005): "Polygyny, Fertility, and Savings," *Journal of Political Economy*, 113, 1341–1371.

Thirumurthy, H., J. Graff Zivin, and M. Goldstein (2008): "The Economic Impact of AIDS Treatment: Labor Supply in Western Kenya," *Journal of Human Resources*, 43, 511–552.

Tompsett, A. (2020): "The Lazarus Drug: The Impact of Antiretroviral Therapy on Economic Growth," *Journal of Development Economics*, 143, 102409.

Transvaal Chamber of Mines (1946): "Map of Southern Africa showing the recruiting systems of the Native Recruiting Corporation and the Witwatersrand Native Labour Association," .

UNAIDS (2004): *Action against AIDS must address epidemic's increasing impact on women, says UN report*.

——— (2020): *UNAIDS Data 2020*, Geneva: Joint United Nations Programme on HIV/AIDS.

——— (2022): "AIDSinfo," .

UNESCO International Institute for Educational Planning, UNESCO Institute for Statistics, Zimbabwe Ministry of Primary and Secondary Education, Zimbabwe Ministry of Higher and Tertiary Education, Science and Technology Development, and Zimbabwe National Statistics Agency (2016): *Report on the Analysis of Public Expenditure on Education in Zimbabwe: Focus on Equity and Efficiency*, Dakar: UNESCO.

Vail, L. (1976): "Mozambique's chartered companies: The rule of the feeble," *Journal of African History*, 17, 389–416.

Viboud, C., L. Simonsen, and G. Chowell (2016): "A Generalized-Growth Model to Characterize the Early Ascending Phase of Infectious Disease," *Epidemics*, 15, 27–37.

Weine, S. M. and A. B. Kashuba (2012): "Labor migration and HIV risk: A systematic review of the literature," *AIDS and Behavior*, 16, 1605–1621.

Weinstein, J. M. (2006): *Inside Rebellion: The Politics of Insurgent Violence*, New York: Cambridge University Press.

Weiss, G., T. Ganz, and L. T. Goodnough (2019): "Anemia of Inflammation," *Blood*, 133, 140–150.

WHO (2002): *Prevention and Control of Schistosomiasis and Soil-Transmitted Helminths: Report of a WHO Expert Committee*, Geneva: World Health Organization.

——— (2020): "Life Tables: Life Tables by Country," .

Wiegand, R. E., W. E. Secor, F. M. Fleming, M. D. French, C. H. King, A. K. Deol, S. P. Montgomery, D. Evans, J. Utzinger, P. Vounatsou, and S. J. de Vlas (2021): "Associations between Infection Intensity Categories and Morbidity Prevalence in School-Age Children Are Much Stronger for S. haematobium than S. mansoni," *PLOS Neglected Tropical Diseases*, 15, e0009444.

Wilson, F. (1972): *Labour in the South African Gold Mines, 1911-1969*, New York: Cambridge University Press.

World Bank (2020): *World Slope Model in Degrees*, The World Bank Group.

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

Young, A. (2005): "The gift of the dying: The tragedy of AIDS and the welfare of future African generations," *Quarterly Journal of Economics*, 120, 423–466.

Zabel, F., B. Putzenlechner, and W. Mauser (2014): "Global agricultural land resources - A high resolution suitability evaluation and its perspectives until 2100 under climate change conditions," *PLOS ONE*, 9, e107522.