The University of Adelaide
Faculty of Professions
School of Economics

ESSAYS ON THE POLITICIAL ECONOMY OF THE HIV/AIDS EPIDEMIC IN SUB-SAHARAN AFRICA

a thesis

by

Pide Lun

submitted in partial fulfilment of
the requirements for the degree of
Doctor of Philosophy

July 16, 2018

Table of Contents

List of Tables

List of Figures

Abstract

Fighting HIV/AIDS remains a priority for sub-Saharan Africa, a region which has been the most severely affected by the disease since the beginning of the epidemic. There is well-documented evidence that the main cause of the rapid spread of HIV in this region is high-risk sexual behavior among the general population. Understanding the motives of this behavior has long been an important research area across a wide range of disciplines. This thesis explores further whether the transmission of HIV in sub-Saharan Africa is associated with factors deeply rooted in culture, history and geography.

First, the thesis provides new evidence that a historical and cultural factor like population genetic diversity can contain the transmission of HIV in sub-Saharan Africa. The thesis examines the impact of population genetic diversity on HIV using trade as an external factor, for there is evidence that trade may propagate HIV infections. Using a difference-in-differences strategy, it is found that in countries that are more genetically homogeneous, HIV spreads more easily. The underlying mechanism is genetic and cultural homophily: in societies that are genetically homogeneous, peoples innate preference for partners who are genetically or culturally similar makes forming sexual relationships easier. Genetic and cultural homophily also allows people to trust each other more. Trust in turn is associated with HIV infections as people express their faithfulness towards their partner through trust by ignoring the perceived risk of contracting HIV.

Second, the thesis shows that the transmission of HIV in sub-Saharan Africa is associated with a geographical feature. The empirical findings suggest that terrain ruggedness, a measure of the uneven topography of an area, has a positive effect on HIV infections. The causal relationship between terrain ruggedness and HIV can

be explained by the following mechanisms. First, terrain ruggedness is negatively associated with historical slave trade in Africa. Rough terrain helped protect those being raided and made it hard to transport slaves from one place to another, thus discouraging slave traders. Secondly, the slave trade matters for the spread of HIV through the conduit of trust. Countries that exported many slaves in the past tend to have societies that are less trusting today; and because low trust is associated with low HIV infections, countries that have high terrain ruggedness could have high rates of HIV infections.

The approach to exploring the underlying causes of HIV transmission in sub-Saharan Africa this thesis has taken up is to question why the HIV epidemic in this region is so uneven. The findings suggest that factors other than socio-economic characteristics have a causal effect on HIV incidence. The historical, cultural and geographical determinants of HIV/AIDS in sub-Saharan Africa suggest that some HIV risk factors are not easily influenced by policies. Moreover, the complexity of the HIV epidemic in sub-Saharan Africa hints that further investigation of the roots of HIV transmission remains a research area worth exploring.

Acknowledgements

I am indebted to many people who contributed to the results of my PhD research and the preparation of this thesis.

First of all, I want to thank my advisor, Dr Nicholas Sim, for his valuable guidance and tremendous support throughout my doctoral research. He helped my analytical skills to flourish and taught me to work independently. Working with Nicholas has been one of the best experiences I have ever had in my academic life.

My gratitude also goes to Dr Terence Cheng, Associate Professor Paul Pezanis-Christou, and Dr Ngoc Thien Anh Pham for their useful advice on how to embark on my research ideas. I am indebted to Dr Firmin Doko Tchatoka and Dr Jacob Wong for their encouragement. I thank Susan Watkins for proofreading some of the chapters and her suggestions to refine my writing.

I am very grateful to the University of Adelaide whose financial support allowed me to stay focused on my studies. I would especially like to thank Associate Professor Mandar Oak whose recommendation opened the door to my admission to the PhD program at the University of Adelaide.

I also enjoyed many conversations with my fellow PhD students in the School of Economics. I thank PhD workshop participants for providing useful comments on my papers. I especially thank Weidong Liang and Dessie Ambew for not only sharing their thoughts and curiosity on my research, but also for their friendship.

Finally, I owe a debt to my wife and daughters for their patience and unconditional love. I dedicate this thesis to them.

Declaration

I, Pide Lun, certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give permission for the digital version of my thesis to be made available on the web, via the Universitys digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Signature: Date: July 16, 2018

Chapter 1

Introduction

This thesis tackles the issue of HIV/AIDS in sub-Saharan Africa, a region which has the highest number of people living with HIV in the world.[1] A broad research question this thesis aims to address is why the spatial distribution of HIV infections in sub-Saharan Africa is so uneven. It is well known that the countries located south and east of the Sahara are most severely affected by HIV/AIDS, while those in west and central sub-Saharan Africa appear to be less affected. Another interesting observation is that the relatively more developed economies in sub-Saharan Africa are not necessarily insulated from HIV.[2] This implies that socio-economic factors alone do not explain why HIV/AIDS is a more serious problem in some countries than in others. In Chapters 3 and 4, we present new evidence suggesting that the differences in the spread of the HIV pandemic in different parts of sub-Saharan Africa have deep historical roots in culture, history and geography.

This thesis begins with a review of literature on HIV in sub-Saharan Africa, with a focus on historical trends in prevalence and incidence rates, its causes and impacts on affected countries, and global responses to date. This chapter documents that although sub-Saharan Africa remains most affected by HIV/AIDS, the trends of HIV epidemic in this region have been declining as evidenced by the falling number of new infections and increased coverage of HIV treatment and care which have contributed

[1] In sub-Saharan Africa, about 25 million people are living with HIV/AIDS and more than 1 million are infected every year.

[2] For example, Botswana is about 10 times richer than Burkina Faso, yet the HIV prevalence rate in Botswana is 15 times greater than that in Burkina Faso.

substantially to the steady decline in the number of AIDS-related deaths. These achievements are due to timely and effective responses by both national governments and the international organizations leading the global combat against HIV/AIDS. Regarding the drivers of HIV, evidence shows that sexual behavior is the main cause of HIV transmission in sub-Saharan Africa. However, factors that influence risky sexual behaviors leading to the spread of HIV are manifold – they include, among others, socio-economic determinants like income and education, cultural norms and values, sexual violence, and migration. Although these factors seem to be associated with the high-profile HIV epidemic in sub-Saharan Africa, the literature is not explicit about whether they also explain the variation in HIV infections in this region. We attempt to uncover the underlying causes of the uneven distribution of HIV infections in sub-Saharan Africa in Chapters 3 and 4.

Chapter 3 examines whether the transmission of HIV in sub-Saharan Africa is rooted in historical-cultural factors such as genetic diversity within populations. Motivated by economic development literature showing that diversity has a negative impact on economic progress, we hypothesize that the transmission of HIV may be constrained by population genetic diversity. The argument unfolds as follows. Firstly, genetic and cultural homophily may facilitate the formation of sexual relationships that are likely to propagate HIV infections. There is evidence that people are sexually attracted to those who share similar attributes to their own. These attributes include physical traits determined by genes and cultural background such as race or religion, which may also be determined by genetic material inherited from ancestors. Secondly, population genetic diversity may have a negative association with trust. In societies that are more diverse, trust is low as people tend to stick to their own group and refuse to cooperate with people from other groups. Trust in turn affects the sexual behavior of the population, for in Africa people tend to underestimate or ignore the potential risks of unprotected sex when they trust their partner. This therefore increases the risk of HIV transmission.

The aim of this chapter is not to demonstrate that population genetic diversity has a direct effect on HIV, but to show that population genetic diversity can restrict the transmission of HIV when caused directly by an external factor such as trade.

To examine this hypothesis empirically, we use a difference-in-differences approach in which trade is employed as the treatment. If genetic diversity makes it harder for HIV to be passed on, trade would then have a weaker effect on HIV incidence in genetically diverse countries (the control group) than in genetically similar countries (the treatment group). To estimate the effect of genetic diversity on HIV through trade, we employ a panel data method known as the common correlated effects (CCE) estimator that simultaneously deals with confounding unobserved country heterogeneity, unobserved macroeconomic factors with potentially heterogeneous effects, and cross-sectional dependence.

Chapter 4 examines the role of physical geography in explaining the spatial variation in HIV infections in sub-Saharan Africa. The initial observation is that sub-Saharan countries with high terrain ruggedness appear to also have high rates of HIV infections. Terrain ruggedness is a geographical measure used to characterize the uneven physical features of an area. Therefore, countries with high terrain ruggedness would have steeply sloping and rocky topography. The association between terrain ruggedness and HIV transmission can be explained by the following mechanisms. Firstly, there is strong evidence that terrain ruggedness has a negative relationship with the slave trade. The African slave trade is a unique period in world history, between 1400 and 1900, during which some 18 million people were enslaved and shipped from Africa to the Middle East, India, Europe and America. This tragic period is marked by the capture and transport of huge numbers of people not to mention the countless deaths of those who died in bondage from brutality. Terrain ruggedness restricted slave exports, for rough topography like cliffs or caves provided protection from raids and subsequent enslavement. Rugged terrain also made it difficult to transport slaves to the nearest coast, and therefore discouraged the capture of slaves. Secondly, there is a relationship between the slave trade and HIV infections through the channel of trust. During the African slave trade, people were not only captured in violent attacks, but were also tricked and sold into slavery by close friends and relatives. Thus people in the community had to stay alert for possible raids or deception by those who might want to sell them into slavery. Living in such a threatening environment created a culture of mistrust which may well have

persisted until the present. Trust, as discussed earlier, may have an effect on HIV transmission as it is associated with risky sexual behavior. To sum up, if terrain ruggedness is associated with lower levels of slave trade, leading to societies that are more trusting, and if trust fuels the spread of HIV, we may expect countries with highly rugged terrain to have a more serious HIV/AIDS problem. In this chapter, we approach our empirical strategy using a panel data method called the within-between estimator. Because terrain ruggedness is a time-invariant variable, we cannot use a fixed effects estimator which employs country fixed effects to control for unobserved permanent differences across countries. Using a random effects method to estimate the impact of terrain ruggedness on HIV may also be problematic if there is a correlation between random effects and other time-variant controls. The within-between estimator addresses the problem of correlated random effects, and at the same time allows us to estimate consistently such time-invariant variables as terrain ruggedness.

This dissertation is organized as follows. Chapter 2 presents a review of literature on HIV/AIDS in sub-Saharan Africa. Chapter 3 examines the impact of population genetic diversity on the transmission of HIV. Chapter 4 further investigates whether the HIV epidemic in sub-Saharan Africa is rooted in a geographical factor measured by terrain ruggedness. Chapter 5 provides concluding remarks.

Chapter 2

A Review of the HIV/AIDS Epidemic in Sub-Saharan Africa

2.1 Introduction

The human immunodeficiency virus (HIV), which causes the acquired immunodeficiency syndrome known as AIDS, is one of the deadliest infectious diseases in history. In sub-Saharan Africa, more than 26 million people have died from HIV/AIDS and the other 25 million are living with the disease (UNAIDS 2017). The annual number of new HIV infections in this region is also high, with more than one million new cases reported in 2016 (UNAIDS 2017). This figure is striking as it accounts for almost two thirds of the global total of new HIV infections. The sheer scale of the HIV epidemic in sub-Saharan Africa not only prompted the global community to focus on controlling HIV/AIDS in this region, but also prompted much research to address the question of why the problem of HIV/AIDS in sub-Saharan Africa is so severe and what can be done to minimize the spread of this disease.

The purpose of this chapter is to review the literature on the HIV pandemic with a focus on sub-Saharan Africa. The literature encompasses a range of disciplines including medicine, epidemiology and economics. The review comprises four main sections. The first describes the history and trends in HIV/AIDS since the start of the epidemic in the mid-to late 1970s. The next section discusses the impact of HIV/AIDS on the affected countries, and the third section brings together all of the evidence pertaining to HIV risk factors and determinants. The final section

gathers information on global responses to HIV/AIDS including HIV prevention and treatment programs. The following summarizes key insights from the literature.

The trends of HIV epidemic in sub-Saharan Africa have been declining, as demonstrated by the falling number of new HIV infections and decreasing number of AIDS-related deaths as a result of effective HIV prevention strategies and increased access to treatment and care. However, during the two decades following the discovery of the disease in the late 1970s, the number of HIV infections increased exponentially resulting in a large number of deaths. The loss of parents to AIDS left millions of children orphaned and in the care of relatives, especially grandparents. Moreover, data on HIV reveals that although sub-Saharan Africa is severely affected by HIV/AIDS, the disease affects the region unevenly. In fact, HIV tends to be much more prevalent in countries in the southern and eastern parts of sub-Saharan Africa than countries in the western and central parts. This stark contrast raises the question of whether there are deep fundamental determinants that explain differences in HIV infections across the region. Motivated by this research question, in the following chapters, we provide new evidence that the HIV epidemic in sub-Saharan Africa is rooted in ingrained factors related to culture, history and physical geography.

Evidence of the negative impacts of HIV/AIDS on affected countries in sub-Saharan Africa is well documented. HIV/AIDS has caused economies to contract and undermined poverty reduction efforts. The disease carries a heavy burden as huge amounts of resources have to be diverted to providing health support and care for HIV-infected populations. Moreover, the adverse health impacts of HIV/AIDS are also enormous. Evidence shows that HIV/AIDS considerably reduces life expectancy and causes fertility to decline. The disease also damages the mental health of both HIV-infected people and their close relatives. HIV patients tend to suffer stigma and stress caused by fear of discrimination and isolation. The relatives of HIV infected people also develop mental health problems as they bemoan the loss of their loved ones and have to shoulder the burden of the impacts on the family that is left behind. The wellbeing of children whose parents are suffering or have died from AIDS is also negatively affected. Those children likely suffer from malnutrition, perform poorly in class and drop out of school early.

In sub-Saharan Africa, risky sexual behavior is found to be the root cause of HIV infections, yet the factors that affect such behavior are complex. Socio-economic factors such as income and education are among those cited. Income shock, for instance, forces young women to engage in transactional sex in which the risk of getting infected is high. Education also determines peoples risk perception about sexually transmitted diseases. For example, the poorly educated may not have access to useful information about how they can protect themselves against such disease. Norms and traditions such as polygyny, and false belief about how HIV is prevented and treated, also explain some risky sexual behaviors. Moreover, there is evidence that sexual violence and abuse against women is associated with higher risk of HIV transmission, while recent literature also suggests that HIV transmission has links with trade, migration and urbanization.

Finally, global responses to HIV/AIDS in sub-Saharan Africa have been outstanding. Several international organizations such as UNAIDS, the Global Fund, WHO and the World Bank have joined forces to lead global actions against HIV/AIDS. Strategies to combat HIV/AIDS focus on both prevention and treatment. HIV prevention programs include blood screening, voluntary counselling and testing (VCT), and education and training especially for high-risk groups such as commercial sex workers, long-distance truck drivers and drug users. The treatment programs involve the use a combination of several medications known as highly active antiretroviral therapy (HAART) to slow the progress of HIV before it develops into full-blown AIDS. These medications have proved very effective and help significantly reduce HIV-related mortality.

2.2 History and Trends

HIV is a deadly disease that progressively destroys the immune system and subsequently develops into AIDS. The virus can be transmitted from an infected person to another through direct contact of body fluids such as blood, semen, vaginal secretions, and breast milk mostly through sexual contact and pregnancy or through syringes and needles used for injecting drugs. There are two types of HIV: HIV-1

is found in most of the HIV-positive population worldwide, while HIV-2 accounts for only a small fraction of the epidemic and is confined mostly to West Africa. Without treatment the HIV infection stage lasts for 10 to 12 years, depending on the individual's general health, before advancing to AIDS. The symptoms of HIV/AIDS were first detected in the late 1970s by doctors in western Africa. Similar cases were later found in Zaire, Congo, Rwanda, Tanzania and Zambia when groups of patients developed various types of fungal diseases unknown to physicians. At that time, little attention was paid by world leaders to the rise of HIV/AIDS despite several activist movements across the globe to demand immediate intervention (UNAIDS 2008).

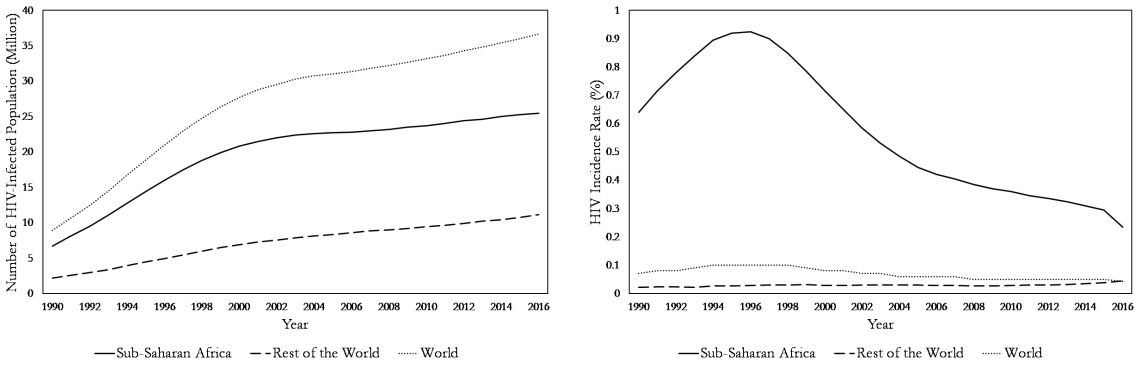
Before 1990, country data on HIV/AIDS was not well organized and published, and only recorded in various antenatal clinics (Oster 2012*b*). The establishment of the Joint United Nations Programme on HIV/AIDS (UNAIDS) in 1995 paved the way for official collection of HIV data; staff were able to trace records back to as early as 1990. By then, HIV/AIDS was continuing to spread across continents, from America to Africa and Asia. According to UNAIDS data,[1] as of 1990, 9 million people worldwide were infected with HIV, 7 million of whom were living in sub-Saharan Africa. In sub-Saharan Africa, like anywhere else, HIV prevalence (the total number of people living with HIV/AIDS) rose steeply following the outbreak in the late 1970s, reaching 21 million in 2000 before the infection rate started to subside. By 2016, sub-Saharan Africa had more than 25 million people living with HIV, accounting for almost 70% of the people infected with HIV worldwide (Figure 2.1, Panel A).

The HIV incidence rate, defined as the annual number of new infections among adults aged 15-49 years, had also risen steeply in sub-Saharan Africa by the end of the 1990s. New cases of HIV among adults aged 15-49 years climbed from 0.60% in 1990 to 0.93% in 1996. The incidence rate has gradually declined since then, largely owing to timely measures taken by national and international communities to control the spread of the virus. Despite these efforts, the average HIV infection rate in sub-Saharan Africa remains relatively high compared to the world average, and to the average in the rest of the world (Figure 2.1, Panel B).

Although statistics suggest that sub-Saharan Africa is severely affected by

[1] See, for example, UNAIDS (2017).

Figure 2.1: HIV Prevalence and Incidence



(A) HIV Prevalence

(B) HIV Incidence

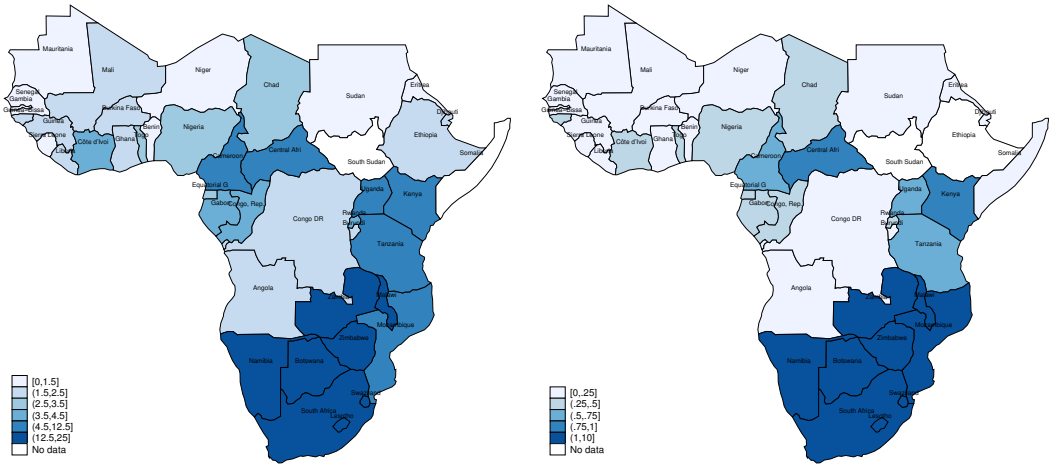
Data source: UNAIDS (2017)

HIV/AIDS, the disease affects the region unevenly. In fact, HIV/AIDS is far more widespread in countries in southern and eastern Africa than in western and central Africa (see Figure 2.2). In 2016, 19 million of the 25.5 million people infected with HIV in sub-Saharan Africa were living in southern and eastern Africa, with countries such as South Africa and Kenya home to the largest numbers of HIV-infected people.[2] The HIV incidence rates among countries in the south and the east are also significantly higher than those in the other parts of sub-Saharan Africa. For example, the average HIV incidence rate in southern and eastern Africa is estimated at around 1.2%, compared to less than 0.5% in western and central Africa. This raises a serious question as to why the variation in HIV infections exists in sub-Saharan Africa. We address this question in Chapters 3 and 4, arguing that historical, cultural and geographical factors specific to certain groups of countries can explain this phenomenon.

HIV prevalence among women and children is at the forefront of the epidemic in sub-Saharan Africa. According to UNAIDS data, in 2016, the prevalence among female adults (aged 15-49) accounted for about 5.5%, compared to some 3.6% for male adults. Young women in sub-Saharan Africa are also at high risk of HIV infection. Recent data suggests that 71% of all adolescents newly infected with HIV are female (UNAIDS 2015). The number of children in sub-Saharan Africa becoming infected

[2]South Africa and Kenya have a combined HIV-positive population of 8.5 million as of 2016.

Figure 2.2: Maps of HIV Prevalence and Incidence



(A) HIV Prevalence (B) HIV Incidence

Data source: UNAIDS (2017)

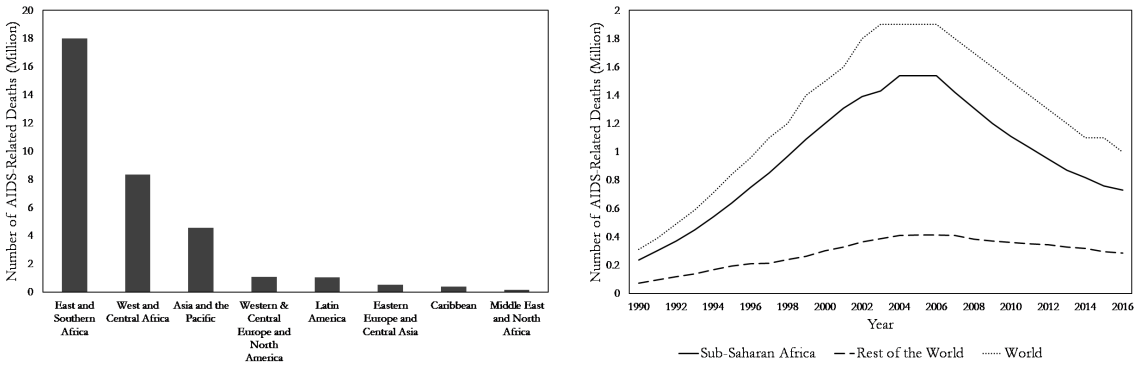
with HIV mostly through their mother is estimated at about 3.2 million or 91% of all HIV-infected children worldwide (UNAIDS 2014).

Today, after wars and conflicts, HIV/AIDS is responsible for the highest number of deaths in sub-Saharan Africa, with more than 26 million people dying from AIDS between 1990 and 2016 (Figure 2.3, Panel (A)). This huge number accounts for more than 75% of the total number of AIDS-related deaths worldwide. As shown in Panel (B), starting from 1990, the number of HIV-related deaths increased exponentially, and peaked in 2004-2006 when the death toll reached 1.5 million. After that, however, the number of deaths plummeted, mainly because of the discovery of antiretroviral therapy (ART) and antiretroviral drugs which suppress the advance of HIV. In 2016, the estimated number deaths from AIDS stood at 0.7 million, more than 50% fewer than the peak number in the mid-2000s, and is expected to decline further.

2.3 Impacts

The adverse impacts of HIV/AIDS on affected countries in sub-Saharan Africa are enormous. On the economy, literature suggests that HIV/AIDS has a negative effect on growth and causes per capita income to decline (WB 1993; Ainsworth and Over 1994). For example, a study by the World Bank estimates that per capita income

Figure 2.3: Number of AIDS-Related Deaths



(A) By Region (1990-2016)

(B) By Region and Year

Data source: UNAIDS (2017)

dropped by around 0.6 percentage points per year in 10 of the sub-Saharan African countries most affected by HIV (WB 1993). This is a consequence of productivity loss in the workforce, and also of productivity decline due to the ill health and high absenteeism of workers infected with HIV (Piot et al. 2001; Fox et al. 2004). Moreover, HIV/AIDS undermines efforts to reduce poverty in sub-Saharan Africa, a region where the majority of people live in poverty. In a poor household, the loss of a family member from AIDS or the poor health of a family member infected with HIV means the loss of household income, exacerbating the already poor economic situation of the household. In addition, HIV/AIDS is also a burden for the affected countries as they have to divert significant resources from productive investments that would boost growth to the health sector in order to provide care and support for HIV/AIDS patients.

The negative impact of HIV/AIDS on health status in heavily affected countries is equally well documented. During the peak of the epidemic, HIV/AIDS plunged adult life expectancy in sub-Saharan African countries to new lows. In Botswana, Lesotho, Swaziland and Zimbabwe, which were hardest hit by HIV/AIDS, adult life expectancies dropped drastically by more than 20 years (WHO 2003). HIV also poses a threat to public health through associated opportunistic infections, notably tuberculosis which is the major cause of death among people living with HIV (WHO 2012). Moreover, evidence of HIV/AIDS deteriorating the mental health of the

population is also well documented. Infected people fear stigma and have a sense of hopelessness, which leads to depression, distress and in the worst cases, suicide (Wingood et al. 2008; Chipimo and Fylkesnes 2009; Visser et al. 2009; Schlebusch and Vawda 2010). The stigma, discrimination and isolation faced by those infected with HIV is caused by fear, while their spouse, parents or close relatives also suffer mental health problems due to caring for or the loss of a loved one and the extra family burden they need to shoulder.

HIV/AIDS also causes fertility to decline. This is explained by the death of a male spouse to HIV or a woman deciding not to have any more children once she knows she is infected with HIV (Lewis et al. 2004). Evidence from medical literature reveals that people infected with HIV, whether due to poor health or symptoms, tend to become less sexually active, and that the virus can stall ovulation or cause miscarriage (Gray et al. 1998; Ross et al. 1999). Other studies suggest that the prevalence of HIV/AIDS alters sexual behaviour through, for instance, delayed marriage, prolonged sexual abstinence, and greater willingness to use barrier methods to protect against HIV (Zaba and Gregson 1998; Lewis et al. 2004).

Finally, HIV/AIDS deals a huge blow to the wellbeing of children who have lost their parents to HIV/AIDS. Economically, the new caregivers, often grandparents who are usually on a limited income, are unable to provide adequate support for those orphans (WHO 2002; Nyasani, Sterberg and Smith 2009). Such resource-constrained settings negatively affect child health as caregivers cannot afford to buy enough food to keep children healthy and nourished (Mishra and Assche 2004; Watts et al. 2007). Mental health issues are also observed among children orphaned by AIDS (Cluver and Orkin 2009). Financial hurdles and poor health consequently disrupt their education in the form of high absenteeism, poor achievement and, in the worst cases, dropping out of school altogether (Andrews, Skinner and Zuma 2006).

2.4 Risk Factors and Determinants

In sub-Saharan Africa, it is evident that risky sexual behavior is the leading cause of HIV infections (Kremer 1996; Oster 2012*a*; Case and Paxson 2013). However, the

factors underlying risky sexual behaviors are complex and manifold. One of these factors is the economic situation. Studies reveal that poverty and income shocks due to, for instance, natural calamities or economic crisis have pushed young women in Africa to get married early to older wealthier men in return for economic stability (Burke, Gong and Jones 2015). Some girls offer transactional sex or in the worst case become sex workers in order to earn income to support their impoverished family (Swidler and Watkins 2007). All of these crisis-coping strategies are associated with a high risk of HIV infection (UNAIDS 2010; Harling et al. 2014). While economic distress makes people adopt sexually risky behaviors and consequently causes HIV to propagate, better economic conditions do not necessarily lower HIV infections. There is evidence that when people earn more, their risk of getting infected also increases as rising income encourages them to have more sexual partners (Oster 2012*b*). The greater the number of sexual partners a person has, the higher the risk of contracting HIV (Aral et al. 1993).

Education also influences sexual behavior, although the relationship between the two is not clear cut. On the one hand, a lack of schooling partly due to economic difficulty may limit people's ability to make rational decisions about safe sex, or their illiteracy denies them access to knowledge and information about how to protect themselves against HIV/AIDS (de Walque 2007, 2009). As such, one effective HIV prevention strategy is to keep young people in school longer. On the other, higher levels of education do not necessarily mean that people will adopt less risky sexual behaviors. High education tends to be correlated with better jobs and good income which, in turn, can increase the demand for sex and hence the number of sexual partners (Glick and Sahn 2008), consequently increasing the risk of contracting HIV. Moreover, formal HIV training itself sometimes fails to materialize. There is evidence that including HIV education in the school curriculum in which girls are advised to not have sex until they are married only ends up encouraging them to get married early and therefore puts them at high risk of infection (Duflo, Dupas and Kremer 2015).

Norms and traditions are also to blame for the spread of HIV/AIDS in sub-Saharan Africa. In some countries, men's extramarital affairs are tolerated and

even formalized. This type of relationship is also known as polygyny, in which men are allowed to have more than one wife or extra-marital relationships without having to conceal them from their official spouse. Having concurrent sex partners is highly risky as people in polygynous relationships are not willing to protect themselves against sexually transmitted diseases including HIV. This is because such relationships are not merely temporary or transactional; they involve love and trust, and could last for a long time (Smith 2007). Sexual protection/condom use in extramarital relationships is viewed as a demonstration of mistrust (Tavory and Swidler 2009; Coma 2014), hence is not favored. Moreover, when it comes to sexual relationships, African women are expected to be submissive. Women dare not discuss their spouse's extramarital affairs, if any, or feel comfortable to negotiate safer sex even if they wish to use protection (Buvé, Bishikwabo-Nsarhaza and Mutangadura 2002).

Violence and sexual abuse also contribute to HIV infections in sub-Saharan Africa. Victims of sexual violence are usually women who are forced to have unwanted and unprotected sex with a partner who is either an alcoholic or drug-addict (Pitpitan et al. 2012). The worst form of sexual violence is rape in which intercourse is mostly unprotected. The high risk of HIV infection through rape is due to genital trauma which increases the risk of HIV transmission, and also because offenders typically have multiple partners and unprotected sex (Klot, Auerbach and Berry 2013).

There is evidence of misconceptions and false beliefs regarding HIV/AIDS in sub-Saharan Africa, which helps explain why some people adopt sexually risky behaviors. It was not that long ago that having HIV was often attributed to witchcraft or sorcery or seen as a punishment from God, and people therefore refused to believe it was purely a medical condition. With this wrong belief, people decided to protect themselves against the disease by seeking supernatural powers rather than choosing HIV prevention methods such as condom use (Tenkorang et al. 2011). Other misconceptions are that having sex with a virgin can cure the disease (de Bruyn 1992) or that HIV can be transmitted through mosquito bites or personal contact (Burgoyne and Drummond 2008). Some people due to their religious beliefs or perhaps poor socio-economic background surrender themselves to God, develop a pessimistic belief that all human beings die anyway, and choose to ignore protection methods even

though they admit that HIV exists (Hess and Mbavu 2010).

Recent literature on economic development provides new evidence on a possible route of HIV transmission in sub-Saharan Africa. For example, Oster (2012*b*) finds that the spread of HIV/AIDS in Africa is rooted in the economic activities of trade. Consistent with historical evidence that trade facilitates the spread of infectious diseases,[3] Oster finds that increased trade is associated with high HIV infections. The channels through which trade affects HIV are trucking and income. Increased trade creates more truck driving jobs. Truck drivers are considered high-risk for they usually seek causal sex along trade routes (Serwadda et al. 1992; Azuonwu, Erhabor and Frank-Peterside 2011). Moreover, it is well documented that trade and income are positively correlated (see *inter alia*, Noguer and Siscart 2005; Manole and Spatareanu 2010); if sex is a normal good, rising income due to increasing trade would encourage individuals to have multiple sexual partners (Oster 2012*b*). As discussed earlier, the risk of HIV infection from having multiple partners is high.

Migration and urbanization also contribute to the spread of HIV. Movement of people away from home in search of employment is often risky when it comes to sexual relationships as migrants are likely to seek casual sex, sometimes without protection (Corno and de Walque 2012). Some migrant workers are susceptible to exploitation and abuse. A case in point is women working in the entertainment sector especially in urban areas (Camlin et al. 2014). Those young women face a high risk of HIV infection because in urban areas where population density is high and the economy is more developed, HIV is also highly prevalent (Dyson 2003; Kelly and van Donk 2009).

Finally, in the medical literature, the transmission of HIV also has links with other sexually transmitted diseases (STIs). Evidence suggests that having an STI increases the risk of HIV infection because genital inflammation facilitates HIV transmission during intercourse (O'Farrell 2001; Galvin and Cohen 2004). Male circumcision on the other hand helps reduce the risk of HIV infection. In some African countries, circumcision is a norm and that plays an important role in HIV prevention (Krantz

[3]For example, the black death in Europe in 14th century was caused by the spread of a bubonic plague originated in Asia and was brought to Europe via trading ships (Pamuk 2007).

and Ahlberg 1995; Hargreave 2010).

2.5 Global Responses

Global responses to HIV/AIDS officially started with the launch in 1986 of the Control Programme on AIDS (CPA)[4] within the World Health Organization (WHO) after several calls by some African countries that HIV/AIDS was a threat. A joint coordinated effort to control the pandemic was proposed following the duplication of HIV/AIDS programs by the United Nations in several countries. This gave rise in 1994 to the Joint United Nations Programme on HIV/AIDS known as UNAIDS. The organization is mandated to lead global action against HIV/AIDS mainly in the areas of prevention, support and treatment (UNAIDS 2008). Besides WHO and UNAIDS, other world-leading institutions have also stepped up to help contribute in the fight against HIV/AIDS. For instance, the World Bank increased loans for HIV-related projects from USD500 million in 1998 to around USD2.7 billion in 2006, mostly for countries in sub-Saharan Africa. In 2001, the ravages of HIV/AIDS beyond Africa prompted the UN Security General to hold a UN Special Session on HIV/AIDS which resulted in the establishment of the multilateral Global Fund to Fight AIDS, Tuberculosis and Malaria in 2002. Shortly afterwards, in 2003, the United States initiated the President's Emergency Plan for AIDS Relief, which allocated USD15 billion for HIV treatment and prevention across 15 countries hardest hit by the epidemic (Merson 2006).

Global actions against HIV/AIDS focus on both HIV prevention and treatment programs, with the former being the most important strategy for minimizing the spread of the disease. To date, numerous HIV prevention projects have been implemented across the globe by both international and local organizations. Firstly, work has been done to ensure that HIV blood screening is readily available so that HIV infection through blood transfusion can eventually be eliminated. Secondly, programs to provide HIV voluntary counselling and testing (VCT) have been put in place everywhere in Africa. VCT is one of the most important HIV prevention

[4]It was renamed Global Programme on AIDS (GPA) in January 1988.

strategies as it provides the means for people to know their HIV status as well as to receive support and care if necessary.

HIV prevention programs have been supplemented by many ongoing projects that provide HIV education worldwide. These include the inclusion of HIV education in school curriculums and peer HIV training among high-risk groups such as commercial sex workers, long-distance truck drivers and drug users. This is to ensure that people are fully aware of the risks, disregard misconceptions about HIV and are able to protect themselves against the virus. HIV education campaigns have come up with many approaches and activities and have even gone on to use technology such as social media and mobile phones to broaden outreach to young sexually active people (UNAIDS 2011).

"Abstinence, be faithful, use a condom" (also known as ABC strategy) remains the key message for HIV education in sub-Saharan Africa, though the effectiveness of such strategy remains debatable. While some country studies contend that the message has contributed to reducing new HIV infections through the adoption of safe sexual behaviors (see *inter alia*, Hogle 2002; USAID 2002; Singh, Darroch and Bankole 2004), others argue that the message is ambiguous in that HIV is not always transmitted through sex (Collins, Coates and Curran 2008). Further, being faithful to a spouse who has extramarital sex does not protect a woman from the risk of contracting HIV (Murphy et al. 2006). Nevertheless, the promotion of condom use has proved to be effective in curbing new HIV infections in sub-Saharan Africa (Myer, Mathews and Little 2002; Ahmed et al. 2011). Such projects include the provision of free condoms in entertainment outlets and brothels, as well as the wider availability of affordable condoms in all possible sale outlets such as pharmacies, hotels, gas stations, and so on.

Another prevention program is to avert new HIV infections among children through mother-to-child transmission during pregnancy, childbirth or breastfeeding. The program has made good progress, mainly due to the invention of antiretroviral (ARV) medicines and therapies and the provision of sufficient support services for HIV-infected mothers. ARV coverage to prevent mother-to-child transmission in sub-Saharan Africa increased from 42% in 2010 to around 70% in 2015 (UNAIDS

2016). Other HIV prevention activities include the treatment of sexually transmitted diseases which facilitate the transmission of HIV (Grosskurth et al. 1995; Oster 2005), the promotion of male circumcision (Weiss and Quigley 2000), and the provision of clean needles and syringes to drug users (WHO 2004).

As for the treatment of HIV/AIDS, it is still a work in progress. There is no cure for the disease, but the discovery of highly active antiretroviral therapy (HAART) in 1996 brought hope to the HIV/AIDS treatment effort. HAART does not cure AIDS nor does it wipe out HIV, yet the therapy combined with the use of antiretroviral drugs helps contain the virus so that it does not progress to AIDS. When HAART was first introduced, only high-income and some middle-income countries could afford its high cost. Only in 2000 did patients in low-income countries in sub-Saharan Africa start to receive HAART treatment following the agreement between UNAIDS and five major pharmaceutical companies as part of the Accelerating Access Initiative (AAI) during the 13th International AIDS Conference in South Africa (Lange and Ananworanich 2014). As of 2015, around 40% of people infected with HIV in sub-Saharan Africa have received ART (UNAIDS 2016), with countries in the east and the south having more access (54%) than those in the west and the east (28%). Looking forward, UNAIDS has already done the groundwork to achieve its ambitious goal to have 90% ART access in all countries by 2020, although today this mission seems somewhat unrealistic.

2.6 Conclusion

To sum up, the prospect of an HIV/AIDS-free sub-Saharan Africa is bright as the world has taken the issue of HIV/AIDS in the region seriously and vowed to eradicate the disease in the very near future. Yet, to achieve this goal there remains a lot of work to be done. HIV prevention approaches need to be more comprehensive and effective, and access to HIV care and support has to be further expanded. HIV research also plays a key role in contributing to this mission, something that this thesis is hoping to offer.

References

Ahmed, Saifuddin, Tom Lutalo, Maria Wawer, David Serwadda, Nelson K. Sewankambo, Fred Nalugoda, Fred Makumbi, Fred Wabwire-Mangen, Noah Kiwanuka, Godfrey Kigozi, Mohamed Kiddugavu and Ron Gray. 2011. "HIV Incidence and Sexually Transmitted Disease Prevalence Associated with Condom Use: A Population Study in Rakai, Uganda." *AIDS* 15(16):2171–2179.

Ainsworth, Martha and Mead Over. 1994. "AIDS and African Development." *World Bank Research Observer* 9(2):203–240.

Andrews, Gail, Donald Skinner and Khangelani Zuma. 2006. "Epidemiology of Health and Vulnerability Among Children Orphaned and Made Vulnerable by HIV/AIDS in Sub-Saharan Africa." *AIDS Care* 18(3):269–276.

Aral, Sevgi O., Mead Over, Lisa Manhart and King K. Holmes. 1993. Sexually Transmitted Infections. In *Disease Control Priorities in Developing Countries*, ed. Dean T. Jamison, Joel G. Breman, Anthony R. Measham, George Alleyne, Mariam Claeson, David B. Evans, Prabhat Jha, Anne Mills and Philip Musgrove. New York: Oxford University Press and the World Bank pp. 311–330.

Azuonwu, Obioma, Osaro Erhabor and Nnenna Frank-Peterside. 2011. "HIV Infection in Long-Distance Truck Drivers in a Low Income Setting in the Niger Delta of Nigeria." *Journal of Community Health* 36(4):583–587.

Burgoyne, Amy D. and Peter D. Drummond. 2008. "Knowledge of HIV and AIDS in Women in Sub-Saharan Africa." *African Journal of Reproductive Health* 12(2):14–31.

Burke, Marshall, Erick Gong and Kelly Jones. 2015. "Income Shocks and HIV in Africa." *Economic Journal* 125(585):1157–1189.

Buvé, Anne, Kizito Bishikwabo-Nsarhaza and Gladys Mutangadura. 2002. "The Spread and Effect of HIV-1 Infection in Sub-Saharan Africa." *Lancet* 359(9322):2011–2017.

Camlin, Carol S., Zachary A. Kwena, Shari L. Dworkin, Craig R. Cohen and Elizabeth A. Bukusi. 2014. "She Mixes Her Business: HIV Transmission and Acquisition Risks Among Female Migrants in Western Kenya." *Social Science and Medicine* 102(2014):146–156.

Case, Anne and Christina Paxson. 2013. "HIV Risk and Adolescent Behaviors in Africa." *American Economic Review* 103(3):433–438.

Chipimo, Peter J. and Knut Fylkesnes. 2009. "Mental Distress in the General Population in Zambia: Impact of HIV and Social Factors." *BMC Public Health* 9(1):298–308.

Cluver, Lucie and Mark Orkin. 2009. "Cumulative Risk and AIDS-Orphanhood: Interactions of Stigma, Bullying and Poverty on Child Mental Health in South Africa." *Social Science and Medicine* 69(8):1186–1193.

Collins, Chris, Thomas J. Coates and James Curran. 2008. "Moving Beyond the Alphabet Soup of HIV Prevention." *AIDS* 22(Supplement 2):S5–S8.

Coma, Julia Cordero. 2014. "HIV Prevention and Marriage: Peer Group Effects on Condom Use Acceptability in Rural Kenya." *Social Science and Medicine* 116(2014):169–177.

Corno, Lucia and Damien de Walque. 2012. "Mines, Migration and HIV/AIDS in Southern Africa." *Journal of African Economies* 21(3):465–498.

de Bruyn, Maria. 1992. "Women and AIDS in Developing Countries." *Social Science and Medicine* 34(3):249–262.

de Walque, Damien. 2007. "How Does the Impact of an HIV/AIDS Information Campaign Vary with Educational Attainment? Evidence from Rural Uganda." *Journal of Development Economics* 84(2):686–714.

de Walque, Damien. 2009. "Does Education Affect HIV Status? Evidence from Five African Countries." *World Bank Economic Review* 23(2):209–233.

Duflo, Esther, Pascaline Dupas and Michael Kremer. 2015. "Education, HIV, and Early Fertility: Experimental Evidence from Kenya." *American Economic Review* 105(9):2757–2597.

Dyson, Tim. 2003. "HIV/AIDS and Urbanization." *Population and Development Review* 29(3):427–442.

Fox, Matthew P., Sydney Rosen, William B. MacLeod, Monique Wasunna, Margaret Bii, Ginamarie Foglia and Jonathon L. Simon. 2004. "The Impact of HIV/AIDS on Labour Productivity in Kenya." *Tropical Medicine and International Health* 9(3):318–324.

Galvin, Shannon R. and Myron S. Cohen. 2004. "The Role of Sexually Transmitted Diseases in HIV Transmission." *Nature Reviews Microbiology* 2(1):33–42.

Glick, Peter J. and David E. Sahn. 2008. "Are Africans Practicing Safer Sex? Evidence from Demographic and Health Surveys for Eight Countries." *Economic Development and Cultural Change* 56(2):397–439.

Gray, Ronald H, Maria J. Wawer, David Serwadda, Nelson Sewankambo, Chuanjun Li, Frederick Wabwire-Mangen, Lynn Paxton, Noah Kiwanuka, Godfrey Kigozi, Joseph Konde-Lule, Thomas C. Quinn, Charlotte A. Gaydos and Denise McNairn. 1998. "Population-Based Study of Fertility in Women with HIV-1 Infection in Uganda." *The Lancet* 351(9096):98–103.

Grosskurth, H., J. Todd, E. Mwijarubi, P. Mayaud, A. Nicoll, G. ka-Gina, J. Newell, D. Mabey, R. Hayes, F. Mosha, K. Senkoro, J. Changalucha, A. Klokke, E. Mwijarubi and K. Mugeye. 1995. "Impact of Improved Treatment of Sexually Transmitted Diseases on HIV Infection in Rural Tanzania: Randomised Controlled Trial." *Lancet* 346(8974):530–536.

Hargreave, Tim. 2010. "Male Circumcision: Cutting into the HIV Epidemic in Africa." *Trends in Urology and Men's Health* 1(2):30–33.

Harling, Guy, Marie-Louise Newell, Frank Tanser, Ichiro Kawachi, SV Subramanian and Till Bärnighausen. 2014. "Do Age-Disparate Relationships Drive HIV Incidence in Young Women? Evidence from a Population Cohort in Rural KwaZulu-Natal, South Africa." *Journal of Acquired Immune Deficiency Syndromes* 66(4):443–451.

Hess, Rosanna F. and Martin Mbavu. 2010. "HIV/AIDS Fatalism, Beliefs and Prevention Indicators in Gabon: Comparisons Between Gabonese and Malians." *African Journal of AIDS Research* 9(2):125–133.

Hogle, Janice A. 2002. *What Happened in Uganda? Declining HIV Prevalence, Behavior Change, and the National Response*. Washington, DC: USAID.

Kelly, Kevin and Mirjam van Donk. 2009. Local-Level Responses to HIV/AIDS in South Africa. In *HIV/AIDS in South Africa 25 Years On: Psychosocial Perspectives*, ed. Poul Rohleder, Leslie Swartz, Seth C. Kalichman and Leickness Chisamu Simbayi. New York: Springer pp. 135–153.

Klot, Jennifer F., Judith D. Auerbach and Miranda R. Berry. 2013. "Sexual Violence and HIV Transmission: Summary Proceedings of a Scientific Research Planning Meeting." *American Journal of Reproductive Immunology* 69(1):5–19.

Krantz, Ingela and Beth Maina Ahlberg. 1995. "Circumcision and HIV." *Lancet* 345(8951):730.

Kremer, Michael. 1996. "Integrating Behavioral Choice into Epidemiological Models of AIDS." *Quarterly Journal of Economics* 111(2):549–573.

Lange, Joep M. A. and Jintanat Ananworanich. 2014. "The Discovery and Development of Antiretroviral Agents." *Antiviral Therapy* 19(Supplement 3):5–14.

Lewis, James J. C., Carine Ronsmans, Alex Ezeh and Simon Gregson. 2004. "The Population Impact of HIV on Fertility in Sub-Saharan Africa." *AIDS* 18(Supplement 2):S35–S43.

Manole, Vlad and Mariana Spatareanu. 2010. "Trade Openness and Income: A Re-examination." *Economics Letters* 106(1):1–3.

Merson, Michael H. 2006. "The HIV-AIDS Pandemic at 25: The Global Response." *New England Journal of Medicine* 354(23):2414–2417.

Mishra, Vinod and Simona Bignami-Van Assche. 2004. Orphans and Vulnerable Children in High HIV-Prevalence Countries in Sub-Saharan Africa. USAID's DHS Analytical Studies No. 14.

Murphy, Elaine M., Margaret E. Greene, Alexandra Mihailovic and Peter Olupot-Olupot. 2006. "Was the "ABC" Approach (Abstinence, Being Faithful, Using Condoms) Responsible for Uganda's Decline in HIV?" *PLoS Medicine* 3(9):1443–1447.

Myer, Landon, Catherine Mathews and Francesca Little. 2002. "Condom Use and Sexual Behaviors Among Individuals Procuring Free Male Condoms in South Africa: A Prospective Study." *Sexually Transmitted Diseases* 29(4):239–241.

Noguer, Marta and Marc Siscart. 2005. "Trade Raises Income: A Precise and Robust Result." *Journal of International Economics* 65(2):447–460.

Nyasani, Evalyne, Erna Sterberg and Helen Smith. 2009. "Fostering Children Affected by AIDS in Richards Bay, South Africa: A Qualitative Study of Grandparents' Experiences." *African Journal of AIDS Research* 8(2):181–192.

O'Farrell, Nigel. 2001. "Targeted Interventions Required Against Genital Ulcers in African Countries Worst Affected by HIV Infection." *Bulletin of the World Health Organization* 79(6):569–577.

Oster, Emily. 2005. "Sexually Transmitted Infections, Sexual Behavior, and the HIV/AIDS Epidemic." *Quarterly Journal of Economics* 120(2):467–515.

Oster, Emily. 2012*a*. "HIV and Sexual Behavior Change: Why Not Africa?" *Journal of Health Economics* 31(1):35–49.

Oster, Emily. 2012*b*. "Routes of Infection: Exports and HIV Incidence in Sub-Saharan Africa." *Journal of the European Economic Association* 10(5):1025–1058.

Pamuk, Şevket. 2007. "The Black Death and the Origins of the 'Great Divergence' Across Europe, 1300-1600." *European Review of Economic History* 11(3):289–317.

Piot, Peter, Michael Bartos, Peter D. Ghys, Neff Walker and Bernhard Schwartlander. 2001. "The Global Impact of HIV/AIDS." *Nature* 410(6831):968–973.

Pitpitan, Eileen V., Seth C. Kalichman, Lisa A. Eaton, Kathleen J. Sikkema, Melissa H. Watt and Donald Skinner. 2012. "Gender-Based Violence and HIV Sexual Risk Behavior: Alcohol Use and Mental Health Problems as Mediators Among Women in Drinking Venues, Cape Town." *Social Science and Medicine* 75(8):1417–1425.

Ross, Amanda, Dilys Morgan, Rosemary Lubega, Lucy M. Carpenter, Billy Mayanja and James A. G. Whitworth. 1999. "Reduced Fertility Associated with HIV: The Contribution of Pre-Existing Subfertility." *AIDS* 13(15):2133–2141.

Schlebusch, Lourens and Naseema Vawda. 2010. "HIV-Infection as a Self-Reported Risk Factor for Attempted Suicide in South Africa." *African Journal of Psychiatry* 13(4):280–283.

Serwadda, David, Maria J. Wawer, Stanley D. Musgrave, Nelson K. Sewankambo, Jonathan E. Kaplan and Ronald H. Gray. 1992. "HIV Risk Factors in Three Geographical Strata of Rural Rakai District, Uganda." *AIDS* 9(6):983–989.

Singh, Susheela, Jacqueline E. Darroch and Akinrinola Bankole. 2004. "A, B and C in Uganda: The Roles of Abstinence, Monogamy and Condom Use in HIV Decline." *Reproductive Health Matters* 12(23):129–131.

Smith, Daniel Jordan. 2007. "Modern Marriage, Men's Extramarital Sex, and HIV Risk in Southeastern Nigeria." *American Journal of Public Health* 97(6):997–1005.

Swidler, Ann and Susan Cotts Watkins. 2007. "Ties of Dependence: AIDS and Transactional Sex in Rural Malawi." *Studies in Family Planning* 38(3):147–162.

Tavory, Iddo and Ann Swidler. 2009. "Condom Semiotics: Meaning and Condom Use in Rural Malawi." *American Sociological Review* 74(2):171–189.

Tenkorang, Eric Y., Stephen O. Gyimah, Eleanor Maticka-Tyndale and Jones Adjei. 2011. "Superstition, Witchcraft and HIV Prevention in Sub-Saharan Africa: The Case of Ghana." *Culture, Health and Sexuality* 13(9):1001–1014.

UNAIDS. 2008. *UNAIDS: The First 10 Years 1996-2007*. Geneva: Joint United Nations Programme on HIV/AIDS.

UNAIDS. 2010. *Report on the Global AIDS Epidemic 2010*. Geneva: Joint United Nations Programme on HIV/AIDS.

UNAIDS. 2011. "Preventing HIV with Social Media and Mobile Phones." http://www.unaids.org/en/resources/presscentre/featurestories/2011/may/20110502sm. (accessed 05 November 2016).

UNAIDS. 2014. *The Gap Report*. Geneva: Joint United Nations Programme on HIV/AIDS.

UNAIDS. 2015. *On the Fast-Track to End AIDS*. Geneva: Joint United Nations Programme on HIV/AIDS.

UNAIDS. 2016. "AIDSinfo." http://aidsinfo.unaids.org/. (accessed 10 January 2017).

UNAIDS. 2017. *UNAIDS Data 2017*. Geneva: Joint United Nations Programme on HIV/AIDS.

USAID. 2002. The "ABCs" of HIV Prevention: Report of a USAID Technical Meeting on Behavior Change Approaches to Primary Prevention of HIV/AIDS. USAID Technical Report.

Visser, Maretha J., Jennifer D. Makin, Alain Vandormael, Kathleen J. Sikkema and Brian W. Forsyth. 2009. "HIV/AIDS Stigma in a South African Community." *AIDS Care* 21(2):197–206.

Watts, Helen, Simon Gregson, Suzue Saito, Ben Lopman, Michael Beasley and Roeland Monasch. 2007. "Poorer Health and Nutritional Outcomes in Orphans and Vulnerable Young Children Not Explained by Greater Exposure to Extreme Poverty in Zimbabwe." *Tropical Medicine and International Health* 12(5):584–593.

WB. 1993. *World Development Report 1993: Investing in Health*. Washington D.C.: World Bank.

Weiss, Helen A. and Maria A. Quigley. 2000. "Male Circumcision and Risk of HIV Infection in Sub-Saharan Africa: A Systematic Review and Meta-Analysis." *AIDS* 14(15):2361–2370.

WHO. 2002. *The Impact of AIDS on the Older People in Africa: Zimbabwe Case Study*. Geneva: World Health Organization.

WHO. 2003. *The World Health Report 2003: Shaping the Future*. Geneva: World Health Organization.

WHO. 2004. Effectiveness of Sterile Needle and Syringe Programming in Reducing HIV/AIDS Among IDUs. Evidence for Action Technical Papers.

WHO. 2012. "TB/HIV Facts 2012-2013." http://www.who.int/hiv/topics/tb/. (accessed 23 December 2016).

Wingood, Gina M., Priscilla Reddy, Shani H. Peterson, Ralph J. DiClemente, Coceka Nogoduka, Nikia Braxton and Anthony D. MBewu. 2008. "HIV Stigma and Mental Health Status Among Women Living with HIV in the Western Cape, South Africa." *South African Journal of Science* 104(5):237–240.

Zaba, Basia and Simon Gregson. 1998. "Measuring the Impact of HIV on Fertility in Africa." *AIDS* 12(Supplement 1):S41–S50.

Chapter 3

Population Diversity, Trust, and the Transmission of HIV/AIDS in Sub-Saharan Africa

Pide Lun Ngoc Pham Nicholas Sim
School of Economics, University of Adelaide

Abstract: We provide new evidence that population genetic diversity can mitigate the transmission of HIV in sub-Saharan Africa. Using a difference-in-differences strategy, we exploit the fact that trade may increase HIV incidence, and demonstrate that HIV transmits more easily in countries that are more genetically homogeneous. Genetically homogeneous societies tend to be culturally homogeneous and more trusting. Thus, our results can be explained by (1) genetic and cultural homophily, as people's innate preference for genetically and culturally similar individuals as partners makes forming relationships easier in genetically and culturally homogeneous societies and (2) trust, as it may reduce one's perceived risk of contracting HIV. Our work highlights the possibility that beyond socio-economic factors, there could exist long-term structural determinants of HIV/AIDS in sub-Saharan Africa that cannot be easily influenced by policies.

Keywords: HIV, sub-Saharan Africa, trade, genetic diversity, cultural diversity, trust.

JEL Classification: I12, O10, O55.

Statement of Authorship

Publication Status	☐ Published ☐ Accepted for Publication ☑ Submitted for Publication ☐ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	Submitted to the Journal of Health Economics in November 2017 as Pide, Lun, Ngoc Pham, and Nicholas Sim. 2017. "Population Diversity, Trust, and the Transmission of HIV/AIDS in Sub-Saharan Africa."

Principal Author

Name of Principal Author (Candidate)	Pide Lun		
Contribution to the Paper	Reviewed literature, collected and analysed data, and wrote the manuscript.		
Overall percentage (%)	65%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	13 January 2018

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Ngoc Pham		
Contribution to the Paper	Analysed data and wrote some parts of the manuscript.		
Signature		Date	13 January 2018

Name of Co-Author	Nicholas Sim		
Contribution to the Paper	Wrote some parts of the manuscript.		
Signature		Date	13 January 2018

Please cut and paste additional co-author panels here as required.

3.1 Introduction

It is well-known that as a region, sub-Saharan Africa suffers the most from the problem of HIV/AIDS.[1] Despite so, there is a large variance in which sub-Saharan African countries (SSAs in short) are affected by this epidemic. One particular nature of this heterogeneity pertains to the fact that countries that suffer more from HIV/AIDS do not necessarily have weaker observable characteristics. For example, Botswana and South Africa have nearly doubled the literacy rate and nearly 10 times the income per capita of Burkina Faso and Ethiopia. However, the HIV/AIDS prevalence rates of Botswana and South Africa are also about 15 times greater.[2] This raises the question of whether there are long-term structural factors, beyond socio-economic characteristics, that could influence the transmission of HIV in the region.

Using panel data on 37 countries from 1985 to 2012, we provide new evidence showing that in the SSAs, the transmission of HIV is associated with within-country population genetic diversity. Genetic diversity refers to the heterogeneity in human physical attributes unique to a certain type of genetic information a person inherits from his or her ancestors. Thus, countries with low genetic diversity have populations that are more physically alike each other.[3] Using the population genetic diversity index constructed by Ashraf and Galor (2013*b*),[4] Figure 3.1 juxtaposes the average yearly HIV incidence rates between 1985 and 2012 and genetic diversity across the SSAs. Here, HIV incidence and population genetic diversity appear to be negatively associated. Specifically, certain countries with higher HIV incidence rates, such

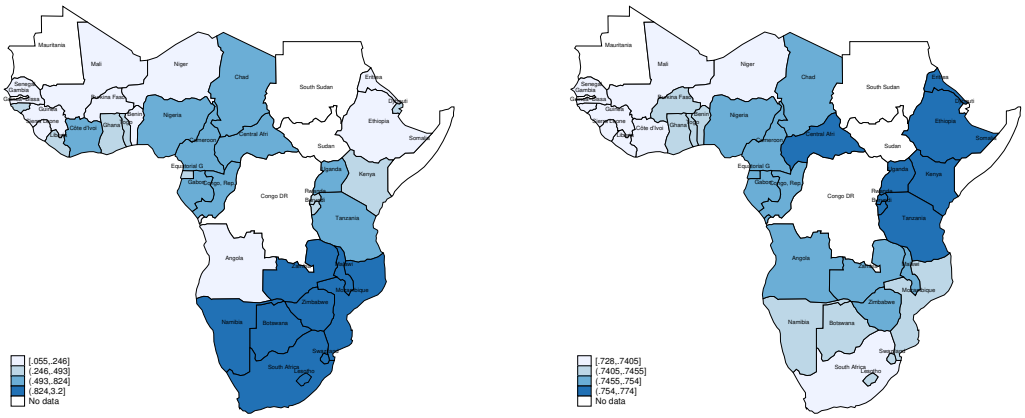
[1] Sub-Saharan Africa accounts for more than two thirds of the world's HIV-positive population (UNAIDS 2014).

[2] In Botswana and South Africa, about one-fifth of the population aged 15-49 are estimated to be HIV positive.

[3] A closely related measure is genetic distance between populations employed in a paper by Spolaore and Wacziarg (2009) and Guiso, Sapienza and Zingales (2009). This indicator uses the US as a base country from which the genetic distance is measured. While genetic distance measures how close the two countries are in terms of genetic traits, genetic diversity on the contrary captures within-country heterogeneity in those traits.

[4] The population genetic diversity index of Ashraf and Galor (2013*b*) is constructed based on genetic diversity as predicted by the migratory distance from East Africa to the rest of the world during the search of new settlements by early mankind. See Section 3.3 for discussion.

Figure 3.1: Geographical Location of HIV Incidence and Genetic Diversity



(A) HIV Incidence (%) (B) Genetic Diversity

as South Africa and Botswana, have populations with lower genetic diversity. By the same token, certain countries like Ethiopia and Kenya, where the problem of HIV/AIDS is less severe, are more genetically diverse.

Population genetic diversity may affect the transmission of HIV for the following reasons. Firstly, the formation of sexual relationships can be motivated by genetic and cultural homophily, as people tend to have intrinsic preference for genetically and culturally similar individuals as partners. For example, in economics, the theory of positive assortative mating predicts that people tend to be more attracted to each other if they both share similar traits and characteristics including physical attributes (Becker 1973, 1974; Greenwood et al. 2014). In psychology and biology, it has also been shown that individuals tend to prefer a partner who looks either like them or their parents (Spuhler 1968; Perrett et al. 2002; Silventoinen et al. 2003; Zietsch et al. 2011). Besides genetic homophily, mutual attraction may also occur among people with similar cultural backgrounds such as race or religion, which are in turn correlated with the genetic information they inherit from their ancestors (Morris 1991; Laumann et al. 1994; Adimora and Schoenbach 2005).

Secondly, population genetic diversity is a source of discontent and mistrust in societies (see, *inter alia*, Alesina and La Ferrara 2005; Guiso, Sapienza and Zingales 2009; Ashraf and Galor 2013*b*). In turn, trust is relevant for how HIV transmits because if a person has a stronger disposition of trust, it may reduce the perceived

level of risk he or she puts on other people or things, such as the risk of contracting the disease.[5] Moreover, since it is not possible to identify HIV positiveness from one's appearance (Bailey et al. 2008), while a person who is HIV positive may refuse to disclose as such for fear of discrimination and the loss of their jobs, health insurance, friends and family (Markel 2003),[6] the belief a person has that his or her partner is HIV-free may influence participation in sex and the nature of the sexual activity (Lear 1995; Mitchell and Horvath 2013). Finally, trust is found to be linked to risky sexual behavior through social norms, where for example, sexual protection through condom use may be construed as an act of mistrust and is therefore eschewed (Pivnick 1993; Appleby, Miller and Rothspan 1999; Alaszewski and Coxon 2009; Coma 2014).

Our main contribution is to show that in the SSAs, something as fundamental as population genetic diversity can influence the transmission of HIV. Our estimation methodology is based on the following empirical design. Suppose for a moment that there exists an external factor from which HIV transmits (i.e. this factor causes HIV incidence to rise). If so, we may employ a difference-in-differences-type approach where this external factor, which causes HIV to propagate, is the treatment; countries with low genetic diversity are the treatment group, and countries with high genetic diversity are the control group. If genetic diversity is *irrelevant* for the spread of HIV, the external factor should affect the HIV incidence rates of countries in both treatment and control groups similarly. However, if genetic diversity makes it harder for HIV to be passed on, this external factor would then influence HIV incidence more weakly in genetically diverse countries (the control group) than in genetically similar countries (the treatment group).

The external factor considered here is trade, as it has been associated with the spread of diseases in the past, such as in the 14^{th} century when the black plague originating from Asia reached the shores of Europe ten years later via trading ships

[5] In lab experiments, trust is found to be negatively associated with lower levels of self-interests and greater levels of commitment and collaborations among individuals (Berg, Dickhaut and McCabe 1995). In finance, low investments in stock markets in some countries could also partly be explained by investors' lack of trust as they believe that firms from which stocks are sold may cheat (Guiso, Sapienza and Zingales 2008).

[6] For example, in a US based study, Stein et al. (1998) found that of sexually active HIV patients in his study, 40% had not disclosed their HIV-positive status to partners whom they had sexual relationship with in the past six months, while only 42% reported that they always used a condom.

and killed a third of Europe's population (Pamuk 2007). For the SSAs, trade is found to have a positive impact on HIV incidence (Kimball 2006; Oster 2012; Lin and Sim 2015).[7] One explanation given in the economics and epidemiology literature is that trade increases trucking; because HIV is more prevalent among truck drivers than the general population is, an increase in trucking resulting from trade may increase the spread of HIV (see, *inter alia*, Serwadda et al. 1992; Azuonwu, Erhabor and Frank-Peterside 2011; Botão et al. 2016).

Pertaining to our difference-in-differences approach, our main explanatory variable is the interaction between 1) an index of genetic diversity, which serves as an indicator of treatment intensity, and 2) trade, which is the treatment itself. To robustly estimate the effect of this interaction term, in a way that deals with confounding unobserved country heterogeneity, unobserved macroeconomic factors with potentially heterogeneous effects, and cross-sectional dependence at the same time, we employ a recently developed though not well-utilized innovation in panel data known as the Common Correlated Effects (CCE) estimator.[8] The CCE approach, developed by Pesaran (2006), enables us to estimate a panel model with interactive fixed effects. Interactive fixed effects contain the additively structured two-way (i.e. country and year) fixed effects as a special case. Therefore, they are more powerful than two-way fixed effects as they can absorb all confounding cross-country permanent differences as country fixed effects can, as well as all confounding macroeconomic (common) factors and country-specific policy shocks that impact countries homogeneously or heterogeneously (Pesaran 2006; Chudik, Pesaran and Tosetti 2011; Chudik and Pesaran 2013), which year fixed effects

[7]For example, Oster (2012) finds that the doubling of exports in the SSAs may increase the HIV incidence rate by 10% to 70%. In addition, Lin and Sim (2015) find that the HIV incidence is 55% higher on average when exports double.

[8]To identify its effect, we have to be mindful of two potential estimation issues. Firstly, trade is not the only macroeconomic factor that can affect HIV incidence. Other factors, such as global economic downturns and the fluctuation of donor countries' currencies, may affect the ability to fund the prevention and treatment of HIV/AIDS (UNAIDS 2009; Jennifer, Wexler and Lief 2016). As these confounding macroeconomic factors may affect the SSAs heterogeneously, they cannot be flushed out simply by using year fixed effects. Secondly, diseases tend to extend across borders. As Figure 3.1 shows, the incidence of HIV/AIDS is not randomly distributed across the SSAs. Instead, it is more concentrated in certain clusters of countries than in others. Such spatial dependence may confound the effect of population genetic diversity on HIV/AIDS incidence that we are trying to identify here.

cannot.[9] Moreover, interactive fixed effects, unlike two-way fixed effects, can also address the potentially confounding issue of cross-sectional dependence that arises from spatial dependence across countries (Chudik, Pesaran and Tosetti 2011). This describes the HIV/AIDS situation in the SSAs, as certain clusters of countries are more severely affected by HIV/AIDS than others. Finally, the consistency of the CCE estimator does not depend on whether the data generating process has an interactive (i.e. cross-sectionally dependent) or two-way fixed effects structure. However, if cross-sectional dependence exists, the two-way fixed effects estimator will be inconsistent.

Our results show that for the SSAs, population genetic diversity may weaken the transmission of HIV. Specifically, while an increase in trade increases the HIV incidence rate on average, as we may expect, this impact is smaller in countries that are more genetically diverse. For instance, comparing the effect of trade per capita on HIV incidence for the "average" country (i.e. whose population genetic diversity is the SSAs' average level), the effect of trade per capita on HIV incidence is 14% *weaker* for a country whose population genetic diversity is one standard deviation *above* the mean. This mitigating effect of genetic diversity is not driven out even as we control for multiple socio-economic characteristics such as income, foreign aid, quality of institutions, education, population density, conflict, incidence of malaria, and weather and commodity price shocks.

Studies have shown that in societies that are more genetically diverse, the diversity of cultures tends to be higher (Ahlerup and Olsson 2012; Ashraf and Galor 2013*a*) and trust tends to be lower (Knack and Keefer 1997; Glaeser et al. 2000). Like the mitigating effect of genetic diversity, we find that the positive association between trade and HIV incidence is weaker in countries with *higher* levels of religious and ethnic diversity and *lower* levels of trust; in other words, in societies that are more religiously and ethnically homogeneous and are more trusting, HIV tends to spread more easily (in this context, when stimulated by trade). However, we also find that the mitigating effect of genetic diversity survives even after partialling out the mitigating

[9]Year fixed effects can only eliminate the confounding effects of macroeconomic factors that affect each cross-sectional unit (e.g. country) *identically*, but not of macroeconomic shocks with cross-sectionally heterogeneous effects.

effects of cultural diversity and trust, which are correlated with genetic diversity itself.[10] This suggests that genetic homophily, along with cultural homophily and trust, are plausible mechanisms behind why in sub-Saharan Africa, HIV transmits more easily in countries with more genetically homogeneous populations.

Over the years, researchers have sought to identify the factors that determine risky sexual behavior, and consequently the spread of HIV/AIDS, particularly in the context of sub-Saharan Africa. What the literature has uncovered include socio-economic factors such as education status (Glick and Sahn 2008; de Walque 2007, 2009; Glick and Sahn 2008; Duflo, Dupas and Kremer 2015), poverty and income shocks (Fenton 2004; Masanjala 2007; Burke, Gong and Jones 2015), norms and traditions such as polygyny (Awusabo-Asare, Anarfi and Agyeman 1993; Bove and Valeggia 2009; Mah and Shelton 2011; Fox 2014), as well as sexual violence (Pitpitan et al. 2012; Jennifer, Wexler and Lief 2016). Our work contributes to the literature by showing that it is possible for a fundamental factor, such as population genetic diversity, to have an effect on the transmission of HIV in the SSAs. This result opens up further discussion on whether the HIV/AIDS problem in the SSAs could be rooted in historical and highly persistent factors such as biology, culture, trust or even geography that are beyond the scope of what policymakers may influence.

The outline of the paper is as follows. Section 2 reviews a number of existing literature connecting population diversity, trust, HIV infections and trade. Section 3 presents the data and summary statistics of key variables. Section 4 discusses the econometric model used to estimate the effect of population genetic diversity on the transmission of HIV caused by trade. Section 5 reports the estimates and discusses robustness results and mechanisms. Section 6 concludes.

3.2 Population Diversity, Trust and HIV/AIDS

In this section, we discuss the relevant evidence that links population genetic diversity to the transmission of HIV. In particular, we discuss how genetic and cultural

[10]The effects of cultural diversity and trust are obtained as the effects of trade interacted with the various measures of cultural diversity and trust.

homophily in sexual preferences, as well as trust, may be linked to the problem of HIV/AIDS. Finally, we provide a brief review on the association between trade and HIV in sub-Saharan Africa. A more technical conceptual framework that explains how population genetic diversity and HIV are linked can be found in Appendix B.

3.2.1 Diversity, Sexual Preference and HIV

The conjecture that diversity and sexual preference are related builds upon the literature on assortative mating and sexual selection. In biometric research, the pioneering work of Pearson (1903) suggests that there is positive assortative mating by physical traits among couples, in that men tend to be attracted to women who have similar appearance to their own, and vice versa. Literature on sexual selection in psychology and biology also supports Pearson's findings that individuals tend to prefer a partner who looks either like them or their parents (Spuhler 1968; Perrett et al. 2002; Bereczkei, Gyuris and Weisfeld 2004; Kocsor et al. 2011; Zietsch et al. 2011). Moreover, evidence of positive assortative mating by non-physical traits among people has been documented as well. These traits include cultural background such as race or religious belief (Adimora and Schoenbach 2005); income and education status (Kalmijn 1994; Laumann et al. 1994); as well as individual personalities (Farky and Muellerms 1978; Le Bon et al. 2013). In economics, the theory of assortative mating is also discussed in the context of a marriage market equilibrium, which predicts that men and women with similar traits match with each other, and in this manner, the total output in the market is maximized (Becker 1973, 1974; Greenwood et al. 2014).

If genetic information determines or influences human traits, and if human traits affect how individuals form relationships, the diversity in genes could influence people's sexual preference and potential number of sexual partners they have. The argument is that similarities in traits like physical appearance and culture among the population could broaden one's sexual networks, thereby increasing the likelihood of sexual engagements or having multiple sexual relationships. Conversely, in populations where human traits are diverse, the innate preference for intra than inter-group sexual relationships creates a natural barrier in the formation of sexual relationships (Laumann and Youm 1999). This may reduce the transmission of HIV,

since the level of sexual activities (e.g. having multiple sexual partners) positively correlates with the risk of contracting HIV (Aral et al. 1993).

3.2.2 Diversity, Trust and HIV

Across several disciplines, there is mounting evidence that trust is negatively associated with diversity. In the economics literature, the early empirical work by Knack and Keefer (1997) shows that high-trust environments coincide with communities that are less ethnically fragmented. This is later corroborated by Glaeser et al. (2000), Alesina and Ferrara (2002) and Guiso, Sapienza and Zingales (2009), who find that trust among people erodes when communities become more diverse in social status, culture, language and political ideology, as distinctive groups tend to stick to their own norms and beliefs and refuse to cooperate with others. More recently, Ashraf and Galor (2013*b*) show that greater differences in human traits, particularly in genetic information, breed low levels of trust. Moreover, diversity also limits civic engagement and community participation, which affects trust. For example, in racially or socially mixed communities, people are reluctant to reach out to unfamiliar groups. As such, their social activities are few and trust declines (Alesina and Ferrara 2000, 2002).

The literature in sociology and political science also lends support to the negative association between trust and diversity. Empirical studies based on North America reveal that higher levels of trust are observed around ethnically homogeneous localities (Hero 2003; Putnam 2007; Soroka, Johnston and Banting 2007; Levine et al. 2014), due to more civic engagement that ties people together (Costa and Kahn 2003). Evidence from cross-country studies also suggests the same (see, *inter alia*, Delhey and Newton 2005; Anderson and Paskeviciute 2006; Kesler and Bloemraad 2010).

In medical research, there has been much discussion pertaining to how trust in conjunction with risk-taking behavior may foster the spread of HIV. For example, it has been found that the element of trust and love undermines the ability of an individual to measure the risk from having sex without protection, while condom use as a risk-coping strategy is viewed as an act of distrust and unfaithfulness (Alaszewski

and Coxon 2009; Tavory and Swidler 2009; Coma 2014; Syvertsen et al. 2015). Moreover, people may choose to ignore the potential consequences of unprotected sex, since trust and love may induce a temptation to take risks. As such, HIV may transmit more easily in more trusting societies.

Extramarital affairs among married men is another conduit through which HIV propagates. Extramarital affairs may involve affection, love and specifically trust (Smith 2007), which reduces the concern of using protection (Mugweni, Pearson and Omar 2015). Thus, married men with extramarital affairs subject their female spouses to the risk of contracting HIV. This problem is further compounded by the fact that in the SSAs, wives are subordinate to their husbands due to traditions and norms, which limits their ability to talk openly about the extramarital affairs of their husbands and more importantly the ability to negotiate with them about whether protection is necessary (Buvé, Bishikwabo-Nsarhaza and Mutangadura 2002). In some cultures, extramarital relationships are made formal and permanent. This practice, known as "polygyny", where men are allowed to have more than one spouse, is found to be associated with the spread of HIV (see *inter alia*, Awusabo-Asare, Anarfi and Agyeman 1993; Bove and Valeggia 2009; Mah and Shelton 2011; Fox 2014).

3.2.3 Trade and HIV

There is well-documented evidence that trade can bring along infectious diseases. One historical example dating to the 14^{th} century is the outbreak of a bubonic plague in Europe, leading to one of the most devastating pandemics in human history known as the Black Death. In this episode, the disease originated in Asia and was carried over to Europe through trading ships, later causing the demise of a third of Europe's population (Pamuk 2007).

Concerning HIV/AIDS, a number of studies have shown that trade may promote the spread of HIV. For example, in the epidemiology literature, Klitsch (1992) observes that people living in trading villages are more susceptible to contracting HIV than those living in non-trading villages, as greater interactions with people

when trade occurs make them more likely engage in multiple sexual activities.[11] Risky sexual activities among people working in trade sector might also be the reason why more trade is correlated with increased HIV infections. Truck drivers, for example, are more likely to be clients of sex workers (Steinbrook 2007). They are considered high risk, as it is not uncommon for them to seek casual sex along trading routes away from home, and once infected, transmit the virus to their spouses at home (Serwadda et al. 1992; Steinbrook 2007; Azuonwu, Erhabor and Frank-Peterside 2011; Botão et al. 2016). As such, an increase in trade, which leads to more trucking, could increase the spread of HIV. Using data from sub-Saharan Africa, Oster (2012) finds that export is the source of HIV infections and that the doubling of exports can cause HIV incidence to increase by 10% to 70%. Similarly, Lin and Sim (2015) find that the number of new HIV infections increases by about 55% when exports double.

3.2.4 Summary

In the literature, there is evidence that the participation in sex is linked to genetic and cultural homophily in sexual preferences (see, *inter alia*, Rushton 1988; Milinski 2006; Kenyon and Colebunders 2013) and trust (see, *inter alia*, Holland et al. 1992; Abbott-Chapman and Denholm 1997; Flood 2003). However, when a society is more genetically diverse, it is likely more culturally diverse and less trusting as well. Because these are barriers in the formation of sexual relationships, the impact of trade on HIV incidence could be weaker in countries with more genetically diverse populations.

3.3 Data and Summary Statistics

Population Diversity Our data are drawn from multiple sources. We extract data on genetic diversity from Ashraf and Galor (2013*b*) who construct a country-level predicted genetic diversity index using historical information on migratory distance

[11] He finds that in Uganda, 62% of men and 32% of women living in trading villages likely had multiple sexual partners within the last five years and that explains why HIV was spread from the trading centers to surrounding rural villages.

from East Africa.[12] The idea is that thousands of years ago, early Homo Sapiens originating from East Africa started to migrate to the rest of the world to find new settlements. However, the farther away a group of human species migrates from the origin, the lower the genetic diversity will be among their population, as they bring with them only a small part of genetic variation of their ancestors (Ashraf and Galor 2013*b*). Empirically, using data from 53 ethnic groups, Ramachandran et al. (2005) show that genetic diversity of these ethnic groups is negatively and strongly associated with their migratory distance from East Africa (the origin of human evolution), which suggests that prehistoric migratory distance from East Africa is a good predictor of population genetic diversity in countries that are present today.

For our empirical analysis, we use Ashraf-Galor's country-level ancestry-adjusted predicted genetic diversity index, which is computed to account for both within- and between-ethnic-group diversity, for it is likely that a country population is made up of multiple ethnic groups from different ancestors due to migration flows since 1500. For example, the ancestry-adjusted predicted genetic diversity of a country populated with two ethnic groups (A and B whose ancestors are from country A and B, respectively) is computed as $\frac{\theta_A \hat{H}^A_{exp}(d_A)+\theta_B \hat{H}^B_{exp}(d_B)}{1-\hat{F}^{AB}_{st}(d_{AB})}$,[13] where $\hat{H}^A_{exp}$ is the predicted genetic diversity by the prehistoric migratory distance between country A to East Africa,[14] θ_A is the share of population of ethnic group A (due to migration) to the total population, and $\hat{F}^{AB}_{st}(d_{AB})$ is the predicted genetic distance between country A and B based on the prehistoric migratory distance between these two countries (Ashraf and Galor 2013*b*).

There are two advantages in using the Ashraf and Galor (2013*b*) index of genetic diversity. Firstly, because it is predicted from the prehistoric migratory distance of a country to East Africa, it can be constructed for as many as 145 countries, and therefore, the countries in our sample. This index is normalized to range from 0

[12] We could consider exploiting the observed genetic diversity index computed by geneticists (see Cavalli-Sforza 2005 for a more detailed discussion of data construction). However, as Ashraf and Galor (2013*b*) explained, the observed genetic diversity index is only available for 53 ethnic groups and limited to 21 countries, which will greatly reduce our sample if we decide to use it.

[13] This equation can be expressed in a recursive form when computing the predicted genetic diversity index of a country comprising more than two ethnic groups (Ashraf and Galor 2013*b*).

[14] Addis Ababa, Ethiopia is the origin.

(complete genetic homogeneity) to 1 (complete genetic heterogeneity), and can be interpreted as the likelihood that two people drawn from a population are genetically different. Secondly, from an econometric modeling perspective, the Ashraf and Galor (2013*b*) genetic diversity index cannot be reverse caused (because it is based on prehistoric information), and this advantage applies here as well.

Besides population genetic diversity, we consider three measures of population *cultural* diversity, which are the diversity in religion, ethnicity and languages. Data on religious and ethnic diversity indices are taken from Montalvo and Reynal-Querol (2005). The indices are defined as $k\sum_{i=1}^{N}\sum_{j\neq i}\pi_i^{\alpha+1}\pi_j$, where π_i and π_j are the proportion of people in religious or ethnic group i and group j, respectively, and k and α are parameters. Montalvo and Reynal-Querol (2005) compute the indices by setting the parameter $k = 4$ and $\alpha = 1$ so that the computed indices meet certain conditions of diversity (see Montalvo and Reynal-Querol 2002) and can be normalized between 0 (no diversity) and 1 (complete diversity). For linguistic diversity, we obtained from Desmet, Ortuño-Ortín and Weber (2009) the Greenberg's index, which measures the population-weighted total linguistic distance between all groups. The index is computed as $\sum_{j=1}^{K}\sum_{k=1}^{K}s_k s_j \tau_{jk}$, where s_k and s_j are the share of population that speaks language k and language j, respectively, and τ_{jk} the continuous measure of language distance between group j and group k.[15] The index is also measured between 0 (no diversity) and 1 (complete diversity).

Trust Our data on trust is taken from the Afrobarometer survey of Round 5 between 2011 and 2013. Afrobarometer is a research network that collects data from many countries in Africa on public attitudes, including information on perception-based trust among their citizens. In Round 5, Afrobarometer designs and collects samples that represent all citizens of voting age (18 years and older) across 35 African countries. Data are weighted to correct for over- or under-sampling and for household size between and within countries.[16]

We extract four different trust indicators including overall trust, trust in relatives,

[15]For example, the language distance between Dutch and Flemish is 0.046, while that between Dutch and Turkish is 1 (Desmet and Weber 2005).

[16]See http://www.afrobarometer.org/about/faqs for a detailed discussion of data construction.

trust in neighbors, and trust in others. The overall trust index is the weighted average of two responses – 1 if most people can be trusted and 0 if one must be very careful – to a survey question that asks if most people in the general population can be trusted or if one must be very careful in dealing with people. The indices on trust in relatives, neighbors, and others (e.g. friends or colleagues) are based on the weighted average of the following four categorical responses – 0 if no trust at all, 1 if just a little, 2 if somewhat, and 3 if a lot – to a survey question that asks how much trusts a person places on his or her relatives, neighbors, and others.

Trade Our trade data is taken from 1) the NBER-UN database for 1985–1994, and 2) the UNCTAD database for 1995–2012.[17] As they are originally expressed in nominal terms, we deflate them by the US urban commodity price index at 2005 US dollars.[18] We compute trade per capita using population data taken from the World Bank's World Development Indicators (WDI).

HIV Incidence Our outcome variable, HIV incidence, is defined as the percentage of new HIV infections among adult population between the age of 15 and 49 in a given year. UNAIDS makes available the data on incidence rates across different countries including the SSAs on their website as well as in their progress reports. Instead of extracting data from UNAIDS, we rely on the dataset used in Oster (2012) in which the HIV incidence rates were recomputed to include those for the years preceding what is available on the UNAIDS website. The main reason for doing so is that the HIV incidence data released by UNAIDS only starts from 1990, but the HIV epidemic goes as far back as the mid-1970s. Oster (2012) calculated the infection rates before 1990 using various country reports on incidence rates at antenatal clinics published by UNAIDS and by adjusting for the number of deaths. The adjusted dataset then spans over 23-year period from 1985 to 2007. We updated the HIV incidence dataset up to the year 2012 using the same method as Oster's. One final point to note is

[17]The NBER-UN trade data spans from 1962 to 2000 (Feenstra et al. 2005), while the UNCTAD data starts from 1995. We compare these two datasets for the overlapping period (i.e. 1995-2000) and find that the difference is insignificant.

[18]See http://data.bls.gov/cgi-bin/surveymost?cu for the data source.

that our dataset on HIV incidence does not perfectly match that of Oster's because, to our best knowledge, the data reported by UNAIDS (which our HIV incidence is calculated from) is constantly being updated. Having said this, the discrepancies between our data and Oster's, if they do arise, are small. For example, for Angola, the average HIV incidence rate between 1985 and 2007 reconstructed using updated data from UNAIDS is estimated to be around 0.18%, which is slightly larger than 0.17% as computed by Oster.

Control Variables For our robustness checks, we employ several time-varying control variables. HIV prevalence, defined as the percentage of individuals found to be infected with HIV, is obtained from Oster (2012). Real Gross Domestic Product, in 2005 US dollars, is sourced from UNCTAD. Polity2, which is the revised Polity score (Polity-2), is obtained from the Polity IV database of Marshall and Jaggers (2009). Polity2 is a common indicator for how democratized a country's political system is. The revised Polity score is computed by combining sub-scores on constraints on the chief executive, the competitiveness of political participation, and the openness and competitiveness of executive recruitment, and ranges from -10 (least democratic) to $+10$ (most democratic). Population density, defined as the total number of population per square kilometer within a country, is retrieved from the World Bank's WDI. Real Official Development Assistance (ODA) per capita, defined as the amount of development aid in US dollars per person deflated with the US urban commodity price index at 2005 US dollars, comes from the UNCTAD database. Primary enrollment rate, defined as the ratio of children of official school age who are enrolled in primary school to the population of the corresponding official school age, is extracted from the World Bank's WDI.

Besides these control variables, we construct control variables related to armed conflict, rainfall, intensity of malaria and international commodity price index. We follow Nunn and Qian (2014) to construct an indicator variable of armed conflict, which is defined as the use of armed force between two parties that has resulted in at least 25 battle deaths in a year. Data on armed conflict is taken from the

UCDP/PRIO Armed Conflict Dataset.[19]

We obtain monthly rainfall data from the World Bank Climate Change Knowledge Portal (CCKP) and construct a weighted average yearly data. According to CCKP, the monthly rainfall dataset is produced by the Climatic Research Unit (CRU) of the University of East Anglia. CRU collects climate station's rainfall records across the world and interpolates rainfall depth (mm) at a $0.5° \times 0.5°$ latitude/longitude grid. It then computes monthly country rainfall series by averaging rainfall depth over all latitude/longitude grid cells covering a country.[20]

Data on the incidence of malaria, defined as the number of new cases of malaria per 100,000 people per year, is extracted from two sources: 1) Montalvo and Reynal-Querol (2007) for the period between 1985 and 1997 and 2) the World Health Organization database for 2000 to 2012. Data on population for each SSA, which is used for constructing the malaria incidence index, comes from the UNCTAD database. We define malaria intensity as $\log(1 + \textit{malaria incidence})$, where 1 is added to malaria incidence to overcome the log transformation problem, since for some countries, there is no information on malaria incidence in certain years.

Lastly, we follow Brückner (2012) to construct a country-specific international commodity price index which is calculated as $\sum_c \theta_{i,c} \log(ComPrice_{c,t})$, where $\log(ComPrice_{c,t})$ is the international price of commodity c[21] in year t (in logs), and $\theta_{i,c}$ is the average (time invariant) value of exports of commodity c over the GDP of country i. We obtain data on annual international commodity prices and the value of commodity exports for the 1985-1994 period from NBER-UN and the 1995-2012 period from UNCTAD. Whenever multiple prices are listed for the same commodity, we use the simple average of these prices as the price of that commodity in our computation. All control variables employed in this paper are lagged by a year, so that we control for their predetermined information.

[19] Available at Uppsala Conflict Data Program/Peace Research Institute website at https://www.prio.org/Data/Armed-Conflict/UCDP-PRIO/.

[20] See http://sdwebx.worldbank.org/climateportal/index.cfm?page=resource#user_guide for Climate Change Knowledge Portal User's Guide and https://crudata.uea.ac.uk/cru/data/hrg/ for high-resolution gridded datasets.

[21] The commodities included in our index are aluminum, bananas, beef, cocoa, coffee, copper, cotton, gold, iron, maize, lead, oil, pepper, rice, rubber, sugar, tea, tobacco, wheat, wood, and zinc.

Table 3.1: Summary Statistics of Main Variables

	Obs.	Mean	Std. Dev.	Min	Max
HIV Indicators (%)					
HIV incidence	1,036	0.774	0.932	0.000	5.152
HIV prevalence	1,036	5.499	6.714	0.001	29.600
Trade Values in 2005 USD					
Trade per capita	986	819.0	2,595	11.780	36,576
Exports per capita	986	481.5	1,837	1.269	25,495
Diversity Measures (0-1)					
Genetic diversity	1,008	0.748	0.011	0.728	0.774
Religious diversity	868	0.811	0.156	0.317	1.000
Ethnic diversity	840	0.544	0.151	0.271	0.843
Linguistic diversity	980	0.187	0.175	0.001	0.591
Trust Indices					
Overall trust	728	0.207	0.154	0.054	0.665
Trust relatives	728	2.411	0.816	1.090	4.140
Trust neighbours	728	1.731	0.641	0.718	3.364
Trust others	728	1.275	0.439	0.483	2.188

Descriptive Statistics In Table 3.1, we provide the basic descriptive statistics for the main variables used in our estimation models. Our dataset contains up to 1,036 observations from 37 SSA countries between 1985 and 2012.[22] The average HIV incidence across countries and years is 0.77%, which means that throughout 1985 to 2012, about 7-8 new cases of HIV per 1,000 per year have been recorded for the SSAs. The standard deviation of HIV incidence (i.e. 0.93%), compared with the mean, is rather large. As emphasized, this is due to the fact that the HIV/AIDS epidemic in the SSAs has substantial regional variations. For instance, HIV/AIDS is especially serious among countries in the southern regions, where the regional average HIV incidence is 1.81%, compared with countries in the central and western regions where the average HIV incidence is less than 0.40%.

The time-invariant genetic diversity in the SSAs has a cross-sectional average of 0.75 and a small standard deviation, with this index varying between 0.72 (Senegal) to 0.77 (Ethiopia).[23] As for cultural diversity, the SSAs, on average, have the highest

[22]Those include Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Central African Republic, Chad, Republic of Congo, Côte d'Ivoire, Djibouti, Equatorial Guinea, Eritrea, Ethiopia, Gabon, Ghana, Guinea, Guinea-Bissau, Kenya, Lesotho, Liberia, Malawi, Mali, Mozambique, Namibia, Niger, Nigeria, Rwanda, Senegal, Sierra Leone, South Africa, Swaziland, Tanzania, Togo, Uganda, Zambia, and Zimbabwe.

[23]While the range of values for genetic diversity in the SSAs is small, it is important to emphasize that the range of values for genetic diversity across the world is somewhat narrow as well, where Uruguay records the lowest genetic diversity index (0.57) and Ethiopia records the highest one (0.77).

diversity of religious belief, compared to that of ethnicity and language. The average value of total trade per capita in the SSAs is about USD 819. The standard deviation of trade per capita is large, owing to the fact that there are large differences in trade volumes between a few relatively wealthy countries like Equatorial Guinea and Gabon and the rest of the SSAs, which are substantially poorer.

3.4 The Model

For country $i = 1, 2, ..., n$ and year $t = 1, 2, ..., T$, we construct a basic unobserved effects panel regression model to estimate the effect of genetic diversity on HIV incidence, which is represented by ability of genetic diversity to potentially mitigate the effect of trade on HIV incidence:

$$H_{it} = \alpha T_{it} + \delta(T_{it} \times G_i) + \boldsymbol{\mu}' \mathbf{X}_{it} + a_i + b_t + \varepsilon_{it}, \tag{3.1}$$

where H_{it} is the HIV incidence; T_{it} is the log of trade per capita; G_i is the genetic diversity; $\mathbf{X}_{it}$ is a vector of year and country level controls to be used in the robustness regressions; and a_i and b_t are the unobserved heterogeneities. The idiosyncratic error term is represented by ε_{it}.

In Eq. (3.1), the variation in T_{it} provides the quasi-natural experiment of generating variations in HIV incidence. The interaction term $T_{it} \times G_i$, whose coefficient is δ, allows for the possibility that population genetic diversity moderates the increase in HIV incidence caused by trade. For a negative δ, an increase in population genetic diversity will weaken the (positive) impact that trade per capita (in log) has on HIV incidence. If δ is statistically insignificant, then there is no evidence that population genetic diversity can mitigate the spread of HIV caused by a rise in trade.

Eq. (3.1) is a two-way fixed effects model. There are two main issues when imposing a two-way fixed effects structure on the unobserved components. Firstly, there could be macroeconomic factors that affect HIV incidence heterogeneously across countries. Some of them include, for instance, global economic downturns

which clearly affect countries heterogeneously, and fluctuations of donor countries' currencies, which may cause funding for HIV/AIDS programs in various recipient countries to decline (UNAIDS 2009; Jennifer, Wexler and Lief 2016). Because the unobserved macroeconomic factors potentially have heterogeneous effects on countries, their confounding effects cannot be dealt with by the use of year fixed effects, as year fixed effects can only deal with confounding macroeconomic shocks that have identical (not heterogeneous) cross-sectional effects.

Secondly, diseases tend to spill across borders. As shown in Figure 3.1, HIV incidence is not randomly distributed across the SSAs. Instead, there are clusters of countries that tend to have higher incidence rates than others. In other words, the severity of the HIV/AIDS problem appears to be spatially dependent. This is evident by the fact that high rates of HIV infection are concentrated along border areas where movement of people from one country to another takes place, as well as along transport routes connecting trade between countries (Morris and Ferguson 2006; UNHCR 2007).

To deal with the confounding influences of unobserved macroeconomic factors with potentially heterogeneous effects, spatial dependence, and the problem of cross-sectional dependence in general, we employ panel regressions with interactive fixed effects and implement the Common Correlated Effects (CCE) estimator of Pesaran (2006). Therefore, instead of Eq. (3.1), we estimate

$$H_{it} = \beta T_{it} + \gamma(T_{it} \times G_i) + \boldsymbol{\psi}' \mathbf{X}_{it} + \boldsymbol{\lambda}_i' \mathbf{F}_t + e_{it}, \tag{3.2}$$

where H_{it}, T_{it}, G_i and $\mathbf{X}_{it}$ represent the same variables as before. Here, $\mathbf{F}_t$ represents the set of unobserved common factors such as global macroeconomic shocks or time trends. When interacted with factor loadings $\boldsymbol{\lambda}_i$ the term $\boldsymbol{\lambda}_i' \mathbf{F}_t$ represents what is known as an interactive fixed effects. Panel estimators that can incorporate interactive fixed effects, such as the CCE estimator, are extremely powerful for dealing with confounding unobserved heterogeneity and macroeconomic factors.

For instance, the unobserved factors contained in $\mathbf{F}_t$ may affect each country differently, as the country factor loadings $\boldsymbol{\lambda}_i$ that $\mathbf{F}_t$ is interacted with are possibly

unique for each country i. Or, the unobserved factors in $\mathbf{F}_t$ may affect each country homogeneously; in which case, the elements in $\boldsymbol{\lambda}_i$ will be identical across i. Because no restrictions about homogeneity or heterogeneity are made on $\boldsymbol{\lambda}_i$, we do not have to impose any priors about whether the unobserved common factors affect countries homogeneously or heterogeneously. Unless we are interested in the common factors themselves, we also do not need to know how many factors are contained in $\mathbf{F}_t$ and what exactly they are, as $\mathbf{F}_t$ represents all relevant unobserved common factors with potentially homogeneous or heterogeneous cross-sectional effects, which will be taken care of by the CCE estimator.

Besides being a powerful approach for dealing with confounding unobserved factors, the CCE estimator enables us to deal with the issue of cross-sectional dependence, including the confounding effects from spatial dependence (Chudik, Pesaran and Tosetti 2011), which appears to be true of the HIV/AIDS problem in the SSAs as discussed earlier.

Finally, the CCE estimator is more robust than the two-way (e.g. country and year) fixed effects estimator in that it subsumes the two-way fixed effects estimator as a special case (Pesaran 2006). For example, if we restrict $\mathbf{F}_t$ and $\boldsymbol{\lambda}_i$ to be

$$\mathbf{F}_t = \begin{bmatrix} 1 \\ \mu_t \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{\lambda}_i = \begin{bmatrix} \mu_i \\ 1 \end{bmatrix},$$

the interactive fixed effects become

$$\boldsymbol{\lambda}_i' \mathbf{F}_t = \mu_i + \mu_t,$$

and Eq. (3.2) boils down to a two-way fixed effects model as a special case. Therefore, if the two-way fixed effects estimator is consistent, the CCE estimator will be consistent. However, if there are common shocks that have heterogeneous effects on countries, or if the problem of HIV/AIDS is cross-sectionally dependent, the two-way fixed effects estimator will be inconsistent, but not the CCE estimator

under certain regularity conditions.[24]

Finally, Pesaran (2006) shows that the covariance matrix can be estimated consistently using the method of Newey and West (1987).[25] As such, for inference under the CCE approach, we report Newey-West standard errors. It should be emphasized that robust standard errors have been suggested by Pesaran (2006) as an alternative. However, we have found both Newey-West and robust standard errors to be very similar, and have no impact on how we conclude in hypothesis tests.

A Comment on Possible Reverse Causality Because our main explanatory variable is the interaction between the log of trade per capita and genetic diversity, we might be concerned that HIV incidence reverse causes trade, or genetic diversity, or both. For example, HIV infections among the workforce may cause declining productivity due to AIDS-related death or sickness. Such weakening of human capital due to HIV/AIDS may then negatively affect economic activities such as trade.

While this is a plausible argument from a long run perspective, in the context of our estimating equation (i.e. Eq. (3.2)), it is unlikely for HIV incidence to reverse cause trade contemporaneously, and thus, for our estimates to be confounded by reverse causality. It should be emphasized that the same point is made by Oster (2012), who argues that reverse causality should not be a concern as the economic impact of new HIV infections is not instantaneous, given that it takes years for an untreated HIV-infected patient to develop severe symptoms or for the virus to progress fully to AIDS.

Likewise, it is not plausible for HIV incidence to have a reverse causal effect on the genetic diversity index. This is because the index is constructed based on prehistoric information, namely, prehistoric migratory paths from East Africa to the rest of the world during the movement of early humans some 150,000 years ago (Ashraf and Galor 2013*b*) (see Section 3).

[24]The conditions for consistency of the CCE estimator are mild (see Assumptions 1-4 of Pesaran, 2006). What is additionally required for the CCE estimator to be consistent over the two-way fixed effects estimator is large N and T. Through a Monte Carlo exercise, Pesaran (2006) shows that the CCE estimator has good small sample properties even when $N = T = 20$, which is easily met in this paper.

[25]See Theorem 1 of Pesaran (2006).

3.5 Empirical Findings

3.5.1 Correlations

We begin by exploring the correlation between our main variables of interest – genetic diversity, trade and HIV incidence – through Figure 3.2, a three-dimensional scatter plot where the genetic diversity index is represented in the x-axis, the log of trade per capita in the y-axis, and the HIV incidence rate in the z-axis. Figure 3.2 shows that as the log of trade per capita is higher, the HIV incidence rate tends to be higher as well. Therefore, trade and HIV incidence appear to be positively correlated, which is not surprising as this positive relationship has already been reported in the literature (Oster 2012; Lin and Sim 2015). By contrast, there appears to be a negative association between HIV incidence and genetic diversity. As shown in Figure 3.2, higher levels of HIV incidence tend to coincide with lower levels of genetic diversity. This negative association is more apparent for the range of log of trade per capita that is smaller than 5.3.

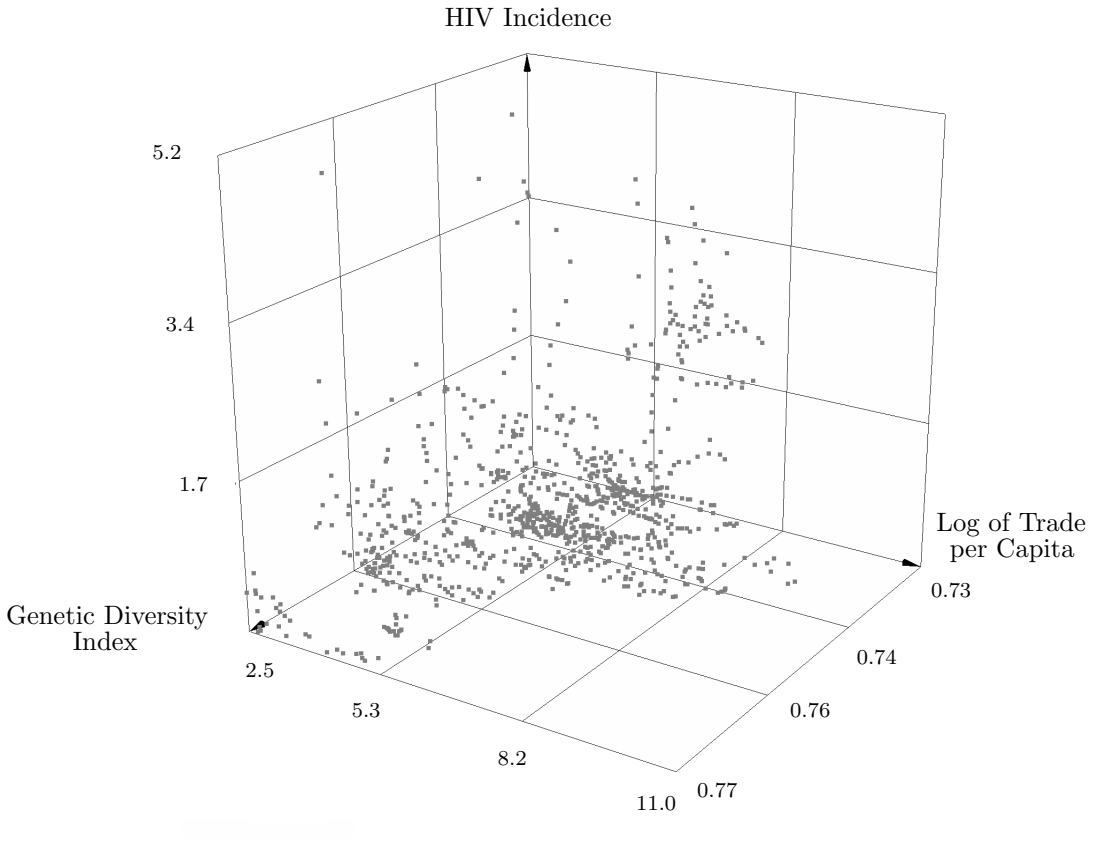
3.5.2 Baseline Estimates

Two-Way Fixed Effects Estimates In Columns (1) and (2) of Table 3.2, we first report the two-way (i.e. country and year) fixed effects estimates of the effect of log of trade per capita and/or its interaction with population diversity on HIV incidence, using the lag of HIV prevalence as an additional control. In Column (1), where the log of trade per capita is included into the regression but not its interaction with genetic diversity, we find that trade has a statistically significant association with HIV incidence, where the doubling of trade per capita is associated with an increase in the HIV incidence rate by about 0.11 percentage point on average.[26] The size of this impact is similar to that reported by Oster (2012), although Oster (2012) uses exports instead of trade.[27]

[26] We consider a 100% increase in trade to avoid unnecessary confusion with the size impact that deals with a few decimal points.

[27] Using Prais-Winsten regression, Oster (2012) finds that the doubling of exports is associated with a 0.07 percentage point increase in HIV incidence.

Figure 3.2: HIV Incidence, Genetic Diversity and Trade



In Column (2), we include the interaction between log of trade per capita and genetic diversity as an additional regressor. We find that the log of trade per capita has a positive effect on HIV incidence as before, but its effect is weaker in countries with higher levels of population genetic diversity. For example, for countries that have the SSAs' average level of genetic diversity (0.748), the doubling of trade per capita is associated with a 0.129 percentage point increase in the HIV incidence rate on average.[28] However, for countries whose population genetic diversity is one standard deviation above the mean (0.759), the doubling of trade per capita is associated with a 0.101 percentage point increase in the HIV incidence rate.[29] Therefore, compared to the effect of trade per capita on HIV incidence for the "average" country (whose genetic diversity level is the SSAs' average), the two-way fixed effects estimate shows that the effect of trade per capita is 22% weaker in countries whose population genetic

[28]This is calculated as $2.030 - (0.748 \times 2.542) = 0.1286$.

[29]This is calculated as $2.030 - (0.759 \times 2.542) = 0.1006$.

Table 3.2: Trade, Genetic Diversity and HIV Incidence: Two-Way Fixed Effects and Interactive Fixed Effects Estimates

	Dependent Variable: HIV Incidence (%)			
	Two-Way Fixed Effects		Interactive Fixed Effects	
	(1)	(2)	(3)	(4)
Log trade per capita	0.109**	2.030*	0.053***	0.652***
	(0.045)	(1.098)	(0.004)	(0.193)
Log trade per capita × genetic diversity		−2.542*		−0.789***
		(1.424)		(0.250)
HIV prevalence$_{t-1}$	0.105***	0.102***	0.120***	0.118***
	(0.007)	(0.008)	(0.004)	(0.004)
P-value CCE	0.015	0.103	–	–
Time trend (general)	Yes	Yes	Yes	Yes
Number of observations	954	927	954	927
Number of countries	37	36	37	36

Note: Clustered robust standard errors (for two-way fixed effects estimates) and Newey-West standard errors (for interactive fixed effects estimates) are reported in the parentheses. The data on HIV incidence are available in 37 SSAs, while 36 countries of which have the data on genetic diversity. Statistical significance at the 10%, 5% and 1% levels are indicated by *, **, and ***, respectively. P-value CCE is used to test a null hypothesis of cross-sectional independence.

diversity is one standard deviation above the mean.

CCE Estimates As discussed, there are good reasons to believe that the SSAs are cross-sectionally dependent.[30] Using the diagnostic test of Pesaran (2004), we test for the presence of cross-sectional dependence against the null of cross-sectional independence, which assumes that the fixed effects have a two-way (additive) structure. In Table 3.2 (under the line *P-value CCE*), the P-values of the test statistics suggest that there is cross-sectional dependence in the HIV incidence rate. For example, in Column (1), the P-value of the cross-sectional dependence test is 1.5%; thus, we reject the null-hypothesis of cross-sectional independence at the 5% level of significance. In Column (2), the P-value increases to 10%, which suggests that some cross-sectional dependence is reduced once we include the interaction between the log of trade per capita and genetic diversity into the model.

Given that cross-sectional dependence can cause the two-way fixed effects

[30]Cross-sectional dependence is likely to arise as the SSAs could be jointly affected by macroeconomic shocks such as global economic downturns and the fluctuation of donor countries' currencies that affect their abilities to fund the prevention and treatment of HIV/AIDS (UNAIDS 2009; Jennifer, Wexler and Lief 2016), or by the spatial nature of the HIV/AIDS problem where certain clusters of countries are more severely affected by it than others. If so, panel regressions that account for cross-sectional dependence, such as the CCE estimator, would be more robust than the two-way fixed effects estimator.

estimator to be inconsistent, a more robust approach is to employ panel regression techniques that incorporate interactive fixed effects, such as the CCE estimator. In Columns (3) and (4), we report the CCE estimates. Qualitatively, the CCE and two-way fixed effects estimates convey the same message that trade has a positive effect on HIV incidence, and that this effect is tempered by higher levels of genetic diversity.

The main difference, however, is that the CCE approach produces smaller estimates of the coefficients on the log of trade per capita and its interaction with genetic diversity. For example, based on the CCE estimates reported in Column (4), we find that a one standard deviation increase in genetic diversity from the mean would reduce the impact that the doubling of trade per capita has on HIV incidence from 0.062 to 0.053 percentage point. In other words, compared with a country with an average level of genetic diversity, a country whose genetic diversity is one standard deviation above the mean encounters an effect of trade on HIV incidence that is 14% weaker. This constraining effect of genetic diversity is smaller than what the two-way fixed effects regression in Column (2) suggests (i.e. 22% based on the same experiment). Overall, while the CCE estimator produces more conservative estimates, our conclusion that genetic diversity can mitigate the transmission of HIV holds regardless of what the fixed effects structure are assumed to be (two-way or interactive).

3.5.3 Sensitivity Analysis

We have found that genetic diversity can mitigate the spread of HIV. We now check for how robust this observation is. To do so, we introduce several new control variables that the literature suggests could be relevant for HIV incidence. If the effect of genetic diversity on HIV incidence reported in the baseline regression is proxying for the effects of these variables, its statistical significance could be driven out once we introduce them as controls into the regression.

Income Recently, there have been studies on how HIV incidence is associated with income in the SSAs. For example, Oster (2012) suggests that income, which is

correlated with trade, has a positive effect on HIV incidence. The author contends that as sex is a normal good, an increase in income would lead to more consumption of sexual activities, and thus, increased risk of HIV infections. However, Burke, Gong and Jones (2015) find that HIV infection rates may increase due to income shocks in Africa's rural areas because households may smooth income in ways (e.g. transactional sex) that contribute to the spread of HIV.

Given that income is strongly associated with trade (Frankel and Romer 1999; Lin and Sim 2013), and it might affect HIV incidence, we check if our conclusion is driven by income. To do so, we introduce the log of real GDP per capita, lagged by one year, as a control variable.[31] The estimates presented in Column (1) of Table 3.3 suggest that income is a relevant determinant of HIV incidence in the SSAs, where the doubling of income per capita in the previous year is associated with a 0.023 percentage point increase in the HIV incidence rate. Despite its statistical significance (at the 1% level), it does not drive out the mitigating effect of genetic diversity on the spread of HIV. In fact, the coefficient on the interaction between the log of trade per capita and genetic diversity still has the same sign, statistical significance, and a similar magnitude as the baseline estimate (Column (4) of Table 3.2).

Foreign Aid Next, we check if our results are driven by the effects of foreign aid. Foreign aid is potentially relevant as it could be channeled into healthcare programs on HIV prevention and treatment (Nunnenkamp and Ohler 2011; Bendavid et al. 2012; UNAIDS 2017). In Column (2) of Table 3.3, we include the log of the Official Development Assistance (ODA) per capita lagged by one year as a control and find it to be statistically insignificant. As such, our baseline estimates are not confounded by foreign aid.

Democracy, Education and Population Density In Columns (3)-(5), we control for institutions, education and population density, respectively. Although it is unclear how institutions are correlated with the problem of HIV/AIDS (Alvaro,

[31] The results (not reported) are similar when contemporaneous values are used instead.

Álvarez-Dardet and Ruiz 2004; Besley and Kudamatsu 2006), there is evidence that democratization may lead to trade liberalization, and thus, more trade (Milner and Kubota 2005). Education and population density are also found to be positively correlated with trade (Keesing and Sherk 1971; Chang and Huang 2014) and associated with the spread of infectious diseases including HIV (Jones et al. 2008).

To control for democracy (Column (3)), we use the Polity2 score from the Polity IV database that captures how democratic a country's institution is. To control for education (Column (4)) and population density (Column (5)), we use the primary enrollment rate and the log of population density (see Section 3.3 for the variables' description). Across Columns (3)-(5), we find that democracy, primary enrollment rate, and population density are statistically significant for HIV incidence at the 1% level. The coefficients on the controls are all positive, and with respect to primary enrollment rate, our result implies rather surprisingly that education is positively correlated with HIV incidence.[32] Most importantly, the effects of trade per capita and its interaction with genetic diversity have the same sign and are similar in size as the estimates in the baseline regression.

Armed Conflict Another potential confounding factor is armed conflict. In the literature, it has been argued that conflict is associated with population diversity, as heterogeneity in the population may give rise to greater differences in ideas and beliefs, and thus, discontent, mistrust, ultimately conflict (Alesina and La Ferrara 2005; Ashraf and Galor 2013*b*). We check if the effect of genetic diversity is capturing the effect of conflict. In the literature, evidence on the relationship between conflict and the problem of HIV/AIDS is mixed.[33] Here, in Column (6) of Table 3.3, we find

[32]Evidence suggests that the better educated tend to have large social networks which make them relatively easy to find sexual partners and hence encourage them to increase the demand for sex, leading to a higher risk of contracting HIV (Glick and Sahn 2008). Similarly, a subsidy and education program to keep girls at school longer and urge them to avoid sex until marriage backfires as it motivates them to get married early and therefore puts themselves at a high risk of HIV infection (Duflo, Dupas and Kremer 2015)

[33]For instance, there is evidence that the HIV/AIDS problem tends to worsen in conflict-affected regions due to increased sexual violence and the disruption of health services (see, for example, USAID 2003; Iqbal and Zorn 2010). However, there is also counter-evidence that conflict slows down the HIV epidemic due to deaths, migration or the destruction of social interaction including sexual relationships (Spiegel et al. 2007).

Table 3.3: Trade, Genetic Diversity and HIV Incidence: Robustness Estimates

	Dependent Variable: HIV Incidence (%)									
Method: CCE	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log trade per capita	0.617***	0.683***	0.833***	0.723**	0.596***	0.924***	0.652***	0.621***	0.577***	1.249***
	(0.186)	(0.230)	(0.251)	(0.284)	(0.203)	(0.184)	(0.205)	(0.192)	(0.152)	(0.320)
Log trade per capita × genetic diversity	−0.764***	−0.829***	−1.029***	−0.894**	−0.708***	−1.159***	−0.790***	−0.741***	−0.676***	−1.464***
	(0.245)	(0.300)	(0.327)	(0.365)	(0.264)	(0.238)	(0.266)	(0.250)	(0.190)	(0.398)
Log GDP per capita $_{t-1}$	0.023***									−0.105***
	(0.004)									(0.011)
Log ODA per capita $_{t-1}$		0.007								0.028***
		(0.011)								(0.007)
Democracy $_{t-1}$			0.006***							0.000
			(0.002)							(0.002)
Primary enrollment $_{t-1}$				0.002***						0.003***
				(0.001)						(0.001)
Log population density $_{t-1}$					0.015***					0.036***
					(0.002)					(0.010)
Conflict $_{t-1}$						−0.085***				−0.078***
						(0.009)				(0.006)
Malaria intensity $_{t-1}$							0.000			−0.003
							(0.003)			(0.003)
Log rainfall $_{t-1}$								−0.068***		−0.150***
								(0.011)		(0.052)
Commodity price $_{t-1}$									−0.174*	−0.033
									(0.092)	(0.132)
HIV prevalence $_{t-1}$	0.118***	0.118***	0.118***	0.117***	0.119***	0.117***	0.118***	0.117***	0.117***	0.110***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
Time trend (general)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Interactive fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	927	918	927	773	927	927	927	927	927	766
Number of countries	36	36	36	36	36	36	36	36	36	36

Note: Newey-West standard errors are reported in the parentheses. The data on HIV incidence are available in 37 SSAs, while 36 countries of which have the data on genetic diversity. Statistical significance at the 10%, 5% and 1% levels are indicated by *, **, and ***, respectively.

that the lag of armed conflict is statistically significant at the 1% level, and negatively associated with HIV incidence. However, after controlling for armed conflict, the mitigating effect of genetic diversity on HIV incidence turns out to be even stronger than what the baseline estimate suggests. For example, based on Column (6), when genetic diversity rises by one standard deviation from the mean, the increase in HIV incidence following the doubling of trade per capita weakens from 0.062 to 0.049 percentage point. In other words, controlling for armed conflict, when genetic diversity increases by one standard deviation from the mean, the marginal effect of trade per capita on HIV incidence weakens by 21%. This reduction is larger than the 14% reported in the baseline regression.

Malaria Incidence Malaria is highly prevalent in the SSAs and could potentially facilitate the transmission of HIV-1 virus (Alemu et al. 2013). In Column (7), we control for a variable on malaria intensity, defined as $\log(1 + malaria\ incidence)$. We find that malaria is not statistically significant for HIV incidence and can therefore be omitted from the regression.

Rainfall and Commodity Price For developing countries that are highly dependent on agriculture, as the SSAs mostly are, rainfall is a key determinant of output, and thus, exports. Given that rainfall may affect trade, it may confound the impact of trade on HIV incidence. Similarly, many SSAs are producers of primary raw materials. Thus, shocks to commodity prices may affect how much the SSAs trade as well.

In Columns (8) and (9), we control for the one-year lag of annual (in log) rainfall and the one-year lag of country-specific commodity price index. We find that rainfall is statistically significant at the 1% level and the commodity price index at the 10% level. However, the coefficients on trade and its interaction with genetic diversity are similar, in terms of sign and size, to the baseline estimates.

All Controls In Column (10), all the control variables are included. Except for foreign aid, malaria incidence, and the commodity price index, the control variables are all statistically significant at the 1% level. Despite this, we find that trade and

its interaction with genetic diversity are statistically significant at the 1% level, and their effects on HIV incidence are similar in magnitude to the baseline CCE estimates (Column (4) of Table 3.2). Importantly, this result also shows that population genetic diversity can mitigate the spread of HIV beyond the effects of socio-economic factors such as income, education, population density, etc.

Summary From Table 3.3, we find that the coefficient on the interaction between log of trade per capita and population genetic diversity is negative and statistically significant at the 1% level throughout the sensitivity analysis. This shows that qualitatively, the mitigating effect of population genetic diversity on the transmission of HIV is robust feature in our regressions.

Further Remarks Firstly, the difference-in-differences approach in our paper uses trade as the treatment. As a robustness check, we have considered using exports in place of trade, and have re-estimated our model with the controls employed in this section.[34] The estimates, presented in Table A1 in Appendix A, show that the mitigating effect of genetic diversity on the spread of HIV is present whether we employ trade or exports as the treatment.

Secondly, instead of using the one-year lag of the variables introduced in this section as controls, we have also considered controlling for their contemporaneous values. Our main conclusion (that genetic diversity can mitigate the transmission of HIV) is robust to using either the lagged or contemporaneous values of these variables as controls.[35]

3.5.4 Possible Mechanisms

Our main result shows that in sub-Saharan Africa, HIV spreads more easily in countries whose population are more genetically homogeneous. One possible mechanism is genetic homophily itself. A second possible mechanism is cultural

[34]The effect of exports on HIV incidence has been considered by (Oster 2012). If the main mechanism for the effect of exports on HIV incidence is trucking, then imports should matter as well. As such, in this paper, we consider trade instead of exports in our estimation approach.

[35]The results are omitted and are available upon request.

homophily, since cultural diversity is correlated with genetic diversity (Ahlerup and Olsson 2012; Ashraf and Galor 2013*a*). As discussed, people have an innate preference for those who are physically or culturally similar to them as partners. Therefore, HIV may transmit more easily in SSAs that are more genetically and culturally homogeneous, as genetic and cultural homophily lowers the barrier among people in the formation of sexual relationships (Milinski 2006; Kenyon and Colebunders 2013).

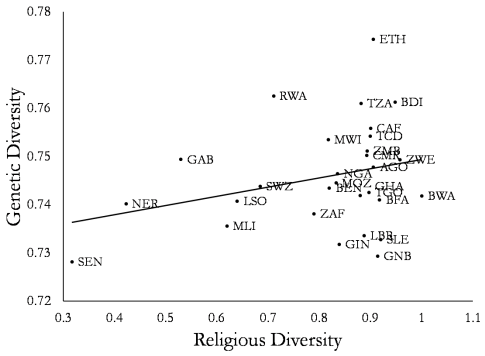
A third possible mechanism is trust. It is well-known that population genetic diversity and trust are negatively associated (Guiso, Sapienza and Zingales 2009; Ashraf and Galor 2013*b*). Since people who are genetically similar tend to trust each other more, this could result in increased risk-taking, especially in risky sexual activities. As such, we cannot rule out that the mitigating effect of genetic diversity on the transmission of HIV is proxying for the effect of the lack of trust (which could be the case if the population is genetically diverse).

To examine if cultural homophily and trust are potential mechanisms, we first visualize how genetic diversity is correlated with cultural diversity and trust. For cultural diversity, we consider three measures, namely, the diversity in religion, ethnicity and languages (see Section 3.3 for their definitions). As for trust, we exploit three measures including overall trust, trust relatives, trust neighbors and trust others (see Section 3.3 for their descriptions). In Figure 3.3, Panels (A)-(C) show that genetic diversity appears to be positively associated with all measures of cultural diversity, although the relationship between genetic diversity and linguistic diversity is rather weak. Panels (D)-(G) further show that genetic diversity is negatively correlated with all measures of trust. As such, the mitigating effect of genetic diversity on the transmission of HIV could be due partly to cultural diversity and trust.

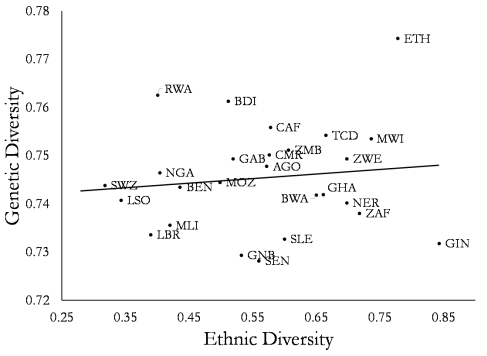
To investigate more formally if cultural diversity can mitigate the spread of HIV, in Table 3.4, we regress the HIV incidence rate on the log of trade per capita and its interaction with religious diversity (Column (1)), ethnic diversity (Column (2)), linguistic diversity (Column (3)), and with all three diversity measures (Column (4)). We also include the full set of control variables that are included in Column (10) of Table 3.3.

As before, trade per capita has a positive and statistically significant effect (at

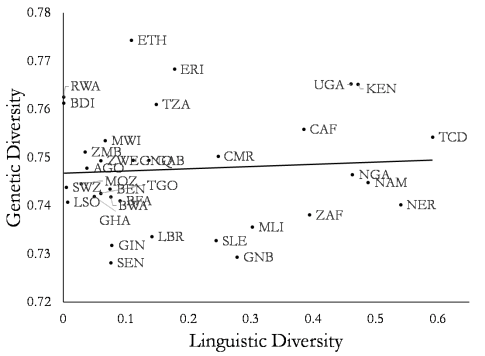
Figure 3.3: Genetic Diversity, Cultural Diversity and Trust



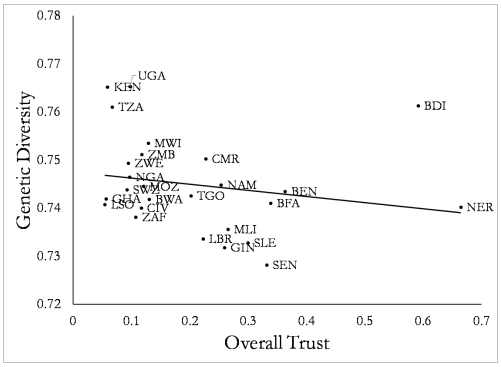
(A) Genetic Diversity and Religious Diversity



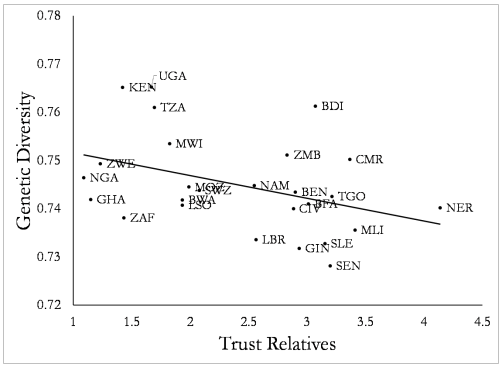
(B) Genetic Diversity and Ethnic Diversity



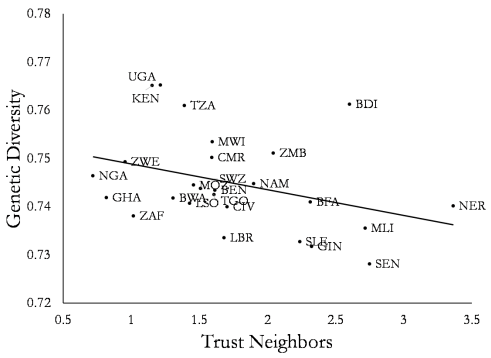
(C) Genetic Diversity and Linguistic Diversity



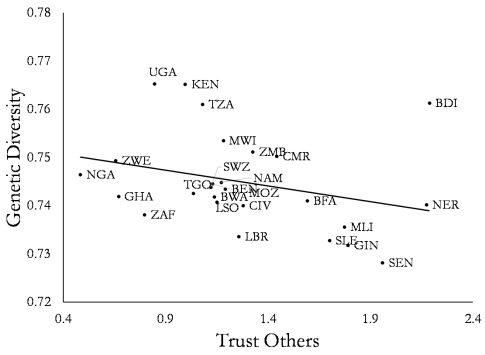
(D) Genetic Diversity and Overall Trust



(E) Genetic Diversity and Trust Relatives



(F) Genetic Diversity and Trust Neighbors



(G) Genetic Diversity and Trust Others

Table 3.4: Trade, Cultural Diversity and HIV Incidence

	Dependent Variable: HIV Incidence (%)			
Method: CCE	(1)	(2)	(3)	(4)
Log trade per capita	0.418***	0.169***	0.103***	0.393***
	(0.065)	(0.015)	(0.028)	(0.053)
Log trade per capita × religious diversity	−0.314***			−0.292***
	(0.074)			(0.064)
Log trade per capita × ethnic diversity		−0.094***		−0.037***
		(0.033)		(0.014)
Log trade per capita × linguistic diversity			0.057	0.058
			(0.039)	(0.040)
HIV prevalence$_{t-1}$	0.107***	0.112***	0.112***	0.107***
	(0.004)	(0.004)	(0.004)	(0.004)
All controls	Yes	Yes	Yes	Yes
Time trend (general)	Yes	Yes	Yes	Yes
Interactive fixed effects	Yes	Yes	Yes	Yes
Number of observations	669	643	741	643
Number of countries	31	30	35	30

Note: Newey-West standard errors are reported in the parentheses. All regressions control for income, foreign aid, democracy, primary enrollment rate, population density, armed conflict, malaria intensity, rainfall, and commodity price. The religious, ethnic and linguistic diversity data are available for 31, 30 and 35 SSA countries, respectively. Statistical significance at the 10%, 5% and 1% levels are indicated by *, **, and ***, respectively.

the 1% level) on HIV incidence. In addition, the interaction of trade per capita (in log) with religious or ethnic diversity has a negative and statistically significant coefficient (Columns (1), (2) and (4) of Table 3.4), but the interaction with linguistic diversity is statistically insignificant (Columns (3) and (4)). This implies that through religious and ethnic diversity, cultural diversity may weaken the impact of trade on HIV incidence. By the same token, it also suggests that cultural homophily, through similarities in one's religion and ethnicity, may foster the spread of HIV.

In Table 3.5, we explore the relationship between trust and HIV incidence. In Column (1), we regress HIV incidence on the log of trade per capita and its interaction with the overall trust measure. We find that at a 1% level of significance, the association between trade and HIV incidence is stronger on average when overall trust is higher.

As an alternative to overall trust, we consider people's trust in their relatives (Column (2)). Like overall trust, we find that the association between trade and HIV incidence is stronger in countries where the trust in relatives is greater. When trust is measured by people's trust in their neighbors (Column (3)) or in other people (Column (4)), we find the impact of trust on HIV incidence to be statistically

Table 3.5: Trade, Population Trust and HIV Incidence

	Dependent Variable: HIV Incidence (%)			
Method: CCE	(1)	(2)	(3)	(4)
Trust Measure:	Overall	Relatives	Neighbors	Others
Log trade per capita	0.157***	0.099**	0.146***	0.162***
	(0.024)	(0.042)	(0.047)	(0.045)
Log trade per capita × trust	0.033***	0.030***	0.010	0.002
	(0.009)	(0.011)	(0.016)	(0.018)
HIV prevalence $_{t-1}$	0.103***	0.101***	0.103***	0.103***
	(0.004)	(0.004)	(0.004)	(0.004)
All controls	Yes	Yes	Yes	Yes
Time trend (general)	Yes	Yes	Yes	Yes
Interactive fixed effects	Yes	Yes	Yes	Yes
Number of observations	575	575	575	575
Number of countries	26	26	26	26

Note: Newey-West standard errors are reported in the parentheses. All regressions control for income, foreign aid, democracy, primary enrollment rate, population density, armed conflict, malaria intensity, rainfall, and commodity price. Although the data on trust taken from the Afrobarometer survey of Round 5 are available for 35 African countries, only 26 of which have HIV incidence data. Statistical significance at the 10%, 5% and 1% levels are indicated by *, **, and ***, respectively.

insignificant. On the whole, if a measure of trust (when interacted with the log of trade per capita) is statistically significant, which is the case for overall trust and the trust in relatives, what we observe is consistent with the notion that trust facilitates the transmission of HIV.

Finally, we explore if population genetic diversity affects HIV incidence independently of its potential correlates – cultural diversity and trust. To do so, we augment our baseline equation with the interaction of trade with all three different aspects of cultural diversity (i.e. religious, ethnic and linguistic) and with a measure of trust based either on overall trust, the trust in relatives, in neighbors, or in others. From Table 3.6, we find that the negative effect of genetic diversity on the transmission of HIV is statistically significant at the 1% level, even after controlling for cultural diversity and trust at the same time.

What do these results suggest about the mechanisms behind the negative effect of genetic diversity on the transmission of HIV? Recall that genetic diversity, which is the diversity in physical traits, is potentially correlated with cultural diversity and trust. Once the effects of cultural diversity and trust are partialled out from the effect of genetic diversity, what should remain in the latter is the effect of the diversity in physical traits on HIV incidence. Given that the cultural diversity and trust (i.e. overall and trust in relatives) are statistically significant (when interacted with

Table 3.6: Genetic Diversity, Cultural Diversity, Trust and HIV Incidence

	Dependent Variable: HIV Incidence (%)			
Method: CCE	(1)	(2)	(3)	(4)
Trust Measure:	Overall	Relatives	Neighbors	Others
Log trade per capita	5.869***	5.233***	5.564***	5.724***
	(1.024)	(1.154)	(1.009)	(0.922)
Log trade per capita × genetic diversity	−6.905***	−6.147***	−6.532***	−6.756***
	(1.415)	(1.562)	(1.346)	(1.232)
Log trade per capita × religious diversity	−0.334***	−0.332***	−0.326***	−0.326***
	(0.142)	(0.136)	(0.123)	(0.128)
Log trade per capita × ethnic diversity	−0.261***	−0.261***	−0.211***	−0.222***
	(0.031)	(0.046)	(0.043)	(0.043)
Log trade per capita × linguistic diversity	0.068***	0.058***	0.065**	0.075**
	(0.023)	(0.022)	(0.030)	(0.035)
Log trade per capita × trust	0.107***	0.036**	0.014	0.025
	(0.018)	(0.015)	(0.019)	(0.023)
HIV prevalence$_{t-1}$	0.085***	0.084***	0.085***	0.085***
	(0.005)	(0.005)	(0.005)	(0.005)
All controls	Yes	Yes	Yes	Yes
Time trend (general)	Yes	Yes	Yes	Yes
Interactive fixed effects	Yes	Yes	Yes	Yes
Number of observations	508	508	508	508
Number of countries	23	23	23	23

Note: Newey-West standard errors are reported in the parentheses. All regressions control for income, foreign aid, democracy, primary enrollment rate, population density, armed conflict, malaria intensity, rainfall, and commodity price. The religious, ethnic and linguistic diversity data are available for 31, 30 and 35 SSA countries, respectively. Although the data on trust taken from the Afrobarometer survey of Round 5 are available for 35 African countries, only 26 of which have data on HIV incidence. Overall, 23 SSAs have the data on genetic diversity, religious, ethnic and linguistic diversity and trust. Statistical significance at the 10%, 5% and 1% levels are indicated by *, **, and ***, respectively.

trade), the statistical significance of genetic diversity suggests that genetic homophily, along with cultural homophily and trust, are all plausible mechanisms behind the negative association between genetic diversity and HIV/AIDS reported in the baseline regression.

3.6 Conclusion

In this paper, we show that a historically and culturally ingrained factor like population genetic diversity can impact the transmission of HIV in sub-Saharan Africa. To do so, we design an empirical strategy that first looks at how an external factor such as trade can cause HIV incidence to rise. Then, we compare this increase in HIV incidence across countries with lower and higher levels of genetic diversity, and find HIV to transmit more easily (i.e. HIV incidence rises more) in countries with

less genetically diverse populations. We also find that genetic homophily, cultural homophily and trust are plausible explanations for why in sub-Saharan Africa, HIV tends to spread more easily in countries that are more genetically homogeneous.

To be clear, the purpose of our paper is to demonstrate that population genetic diversity can restrict the transmission of HIV when caused directly by an external factor such as trade. It does not argue that genetic diversity is itself a direct cause of HIV in the SSAs. Nonetheless, the robustness of our findings (that genetic diversity affects the transmission of HIV) suggests that this might be worthwhile to pursue further, and that by focusing on deep fundamentals such as population genetic diversity, and to some extent, population cultural diversity and trust, as possible direct determinants of HIV/AIDS, we might gain further insights into why certain countries in sub-Saharan Africa are more severely affected by the HIV/AIDS epidemic than others.

References

Abbott-Chapman, Joan and Carey Denholm. 1997. "Adolescent Risk Taking and the Romantic Ethic: HIV/AIDS Awareness Among Year 11 and 12 Students." *Australian and New Zealand Journal of Sociology* 33(3):306–321.

Adimora, Adaora A. and Victor J. Schoenbach. 2005. "Social Context, Sexual Networks, and Racial Disparities in Rates of Sexually Transmitted Infections." *Journal of Infectious Diseases* 191(Supplement1):S115–S122.

Ahlerup, Pelle and Ola Olsson. 2012. "The Roots of Ethnic Diversity." *Journal of Economic Growth* 17(2):71–102.

Alaszewski, Andy and Kirstie Coxon. 2009. "Uncertainty in Everyday Life: Risk, Worry and Trust." *Health, Risk and Society* 11(3):201–207.

Alemu, Abebe, Yitayal Shiferaw, Zelalem Addis, Biniam Mathewos and Wubet Birhan. 2013. "Effect of Malaria on HIV/AIDS Transmission and Progression." *Parasites and Vectors* 6(18):1–8.

Alesina, Alberto and Eliana La Ferrara. 2000. "Participation in Heterogeneous Communities." *Quarterly Journal of Economics* 115(3):847–904.

Alesina, Alberto and Eliana La Ferrara. 2002. "Who Trusts Others?" *Journal of Public Economics* 85(2):207–234.

Alesina, Alberto and Eliana La Ferrara. 2005. "Ethnic Diversity and Economic Performance." *Journal of Economic Literature* 43(3):762–800.

Alvaro, Franco, Carlos Álvarez-Dardet and Maria Ruiz. 2004. "Effect of Democracy on Health: Ecological Study." *British Medical Journal* 329(7480):1421–1423.

Anderson, Christopher J. and Aida Paskeviciute. 2006. "How Ethnic and Linguistic Heterogeneity Influence the Prospects for Civil Society: A Comparative Study of Citizenship Behavior." *Journal of Politics* 68(4):783–802.

Appleby, Paul R., Lynn C. Miller and Sadina Rothspan. 1999. "The Paradox of Trust for Male Couples: When Risking is a Part of Loving." *Personal Relationships* 6(1):81–93.

Aral, Sevgi O., Mead Over, Lisa Manhart and King K. Holmes. 1993. Sexually Transmitted Infections. In *Disease Control Priorities in Developing Countries*, ed. Dean T. Jamison, Joel G. Breman, Anthony R. Measham, George Alleyne, Mariam Claeson, David B. Evans, Prabhat Jha, Anne Mills and Philip Musgrove. New York: Oxford University Press and the World Bank pp. 311–330.

Ashraf, Quamrul and Oded Galor. 2013*a*. "Genetic Diversity and the Origins of Cultural Fragmentation." *American Economic Review* 103(3):528–533.

Ashraf, Quamrul and Oded Galor. 2013*b*. "The 'Out of Africa' Hypothesis, Human Genetic Diversity, and Comparative Economic Development." *American Economic Review* 103(1):1–46.

Awusabo-Asare, Kofi, John K. Anarfi and D.K. Agyeman. 1993. "Women's Control Over Their Sexuality and the Spread of STDs and HIV/AIDS in Ghana." *Health Transition Review* 3(Supplementary):69–84.

Azuonwu, Obioma, Osaro Erhabor and Nnenna Frank-Peterside. 2011. "HIV Infection in Long-Distance Truck Drivers in a Low Income Setting in the Niger Delta of Nigeria." *Journal of Community Health* 36(4):583–587.

Bailey, Justin R., Karen O'Connell, Hung-Chih Yang, Jie Han, Yefei Xu, Benjamin Jilek, Thomas M. Williams, Robert F. Ray, Stuart C. Siliciano and Joel N. Blankson. 2008. "Transmission of Human Immunodeficiency Virus Type 1 from a Patient Who Developed AIDS to an Elite Suppressor." *Journal of Virology* 82(15):7395–7410.

Becker, Gary S. 1973. "A Theory of Marriage: Part I." *Journal of Political Economy* 81(4):813–846.

Becker, Gary S. 1974. "A Theory of Marriage: Part II." *Journal of Political Economy* 82(2):S11–S26.

Bendavid, Eran, Charles B. Holmes, Jayanta Bhattacharya and Grant Miller. 2012. "HIV Development Assistance and Adult Mortality in Africa." *JAMA* 307(19):2060–2067.

Bereczkei, Tamas, Petra Gyuris and Glenn E. Weisfeld. 2004. "Sexual Imprinting in Human Mate Choice." *Proceedings: Biological Sciences* 271(1544):1129–1134.

Berg, Joyce, John Dickhaut and Kevin McCabe. 1995. "Trust, Reciprocity, and Social History." *Games and Economic Behavior* 10(1):122–142.

Besley, Timothy and Masayuki Kudamatsu. 2006. "Health and Democracy." *American Economic Review* 96(2):313–318.

Botão, Carlos, Roberta Z. Horth, Heidi Frank, Beverley Cummings, Celso Inguane, Isabel Sathane, Willi McFarland, Raymond H. Fisher and Peter W. Young. 2016. "Prevalence of HIV and Associated Risk Factors Among Long Distance Truck Drivers in Inchope, Mozambique, 2012." *AIDS and Behavior* 20(4):811–820.

Bove, Riley and Claudia Valeggia. 2009. "Polygyny and Women's Health in Sub-Saharan Africa." *Social Science and Medicine* 68(1):21–29.

Brückner, Markus. 2012. "Economic Growth, Size of Agricultural Sector, and Urbanization in Africa." *Journal of Urban Economics* 71(1):26–36.

Burke, Marshall, Erick Gong and Kelly Jones. 2015. "Income Shocks and HIV in Africa." *Economic Journal* 125(585):1157–1189.

Buvé, Anne, Kizito Bishikwabo-Nsarhaza and Gladys Mutangadura. 2002. "The Spread and Effect of HIV-1 Infection in Sub-Saharan Africa." *Lancet* 359(9322):2011–2017.

Cavalli-Sforza, Luca L. 2005. "The Human Genome Diversity Project: Past, Present and Future." *Nature Reviews Genetics* 6(4):333–340.

CDC (Centers for Disease Control and Prevention). 2016. "Gay and Bisexual Men's Health." http://www.cdc.gov/msmhealth/HIV.htm. (Accessed February 2017).

Chang, Pao-Li and Fali Huang. 2014. "Trade and Divergence in Education System." *International Economic Review* 55(4):1251–1280.

Chudik, Alexander and Hashem M. Pesaran. 2013. Large Panel Data Models with Cross-Sectional Dependence: A Survey. CESifo Working Paper Series No. 4371.

Chudik, Alexander, Hashem M. Pesaran and Elisa Tosetti. 2011. "Weak and Strong Cross-Section Dependence and Estimation of Large Panels." *Econometrics Journal* 14(1):C45–C90.

Coma, Julia C. 2014. "HIV Prevention and Marriage: Peer Group Effects on Condom Use Acceptability in Rural Kenya." *Social Science and Medicine* 116(2014):169–177.

Costa, Dora L. and Matthew E. Kahn. 2003. "Civic Engagement and Community Heterogeneity: An Economist's Perspective." *Perspectives on Politics* 1(1):103–111.

de Walque, Damien. 2007. "How Does the Impact of an HIV/AIDS Information Campaign Vary with Educational Attainment? Evidence from Rural Uganda." *Journal of Development Economics* 84(2):686–714.

de Walque, Damien. 2009. "Does Education Affect HIV Status? Evidence from Five African Countries." *World Bank Economic Review* 23(2):209–233.

Delhey, Jan and Kenneth Newton. 2005. "Predicting Cross-National Levels of Social Trust: Global Pattern or Nordic Exceptionalism?" *European Sociological Review* 21(4):311–327.

Desmet, Klaus, Ignacio Ortuño-Ortn and Shlomo Weber. 2005. Peripheral Diversity and Redistribution. CEPR Discussion Paper No. 5112.

Desmet, Klaus, Ignacio Ortuño-Ortín and Shlomo Weber. 2009. "Linguistic Diversity and Redistribution." *Journal of the European Economic Association* 7(6):1291–1318.

Duflo, Esther, Pascaline Dupas and Michael Kremer. 2015. "Education, HIV, and Early Fertility: Experimental Evidence from Kenya." *American Economic Review* 105(9):2757–2597.

Farky, Frank H. and Carol B. Muellerms. 1978. "Arousal, Personality, and Assortative Mating in Marriage: Generalizability and Cross-Cultural Factors." *Journal of Sex and Marital Therapy* 4(1):50–53.

Feenstra, Robert, Robert Lipsey, Haiyan Deng, Alyson Ma and Hengyong Mo. 2005. "World Trade Flows, 1962-2000." NBER Working Paper No. 11040.

Fenton, Lynda. 2004. "Preventing HIV/AIDS Through Poverty Reduction: The Only Sustainable Solution?" *Lancet* 364(9440):1186–1187.

Flood, Michael. 2003. "Lust, Trust and Latex: Why Young Heterosexual Men Do Not Use Condoms." *Culture, Health and Sexuality* 5(4):353–369.

Fox, Ashley M. 2014. "Marital Concurrency and HIV Risk in 16 African Countries." *AIDS and Behavior* 18(4):791–800.

Frankel, Jeffrey and David Romer. 1999. "Does Trade Cause Growth." *American Economic Review* 89(3):379–399.

Garnett, Geoffrey P. 2002. "An Introduction to Mathematical Models in Sexually Transmitted Disease Epidemiology." *Sexually Transmitted Infections* 78(1):7–12.

Glaeser, Edward L., David I. Laibson, Jose A. Scheinkman and Christine L. Soutter. 2000. "Measuring Trust." *Quarterly Journal of Economics* 115(3):811–846.

Glick, Peter J. and David E. Sahn. 2008. "Are Africans Practicing Safer Sex? Evidence from Demographic and Health Surveys for Eight Countries." *Economic Development and Cultural Change* 56(2):397–439.

Greenwood, Jeremy, Nezih Guner, Georgi Kocharkov and Cezar Santos. 2014. "Marry Your Like: Assortative Mating and Income Inequality." *American Economic Review* 104(5):348–353.

Greenwood, Jeremy, Philipp Kircher, Cezar Santos and Michelè Tertilt. 2013. "An Equilibrium Model of the African HIV/AIDS Epidemic." NBER Working paper No. 18953.

Guiso, Luigi, Paola Sapienza and Luigi Zingales. 2008. "Trusting the Stock Market." *Journal of Finance* 63(6):2557–2600.

Guiso, Luigi, Paola Sapienza and Luigi Zingales. 2009. "Cultural Biases in Economic Exchange?" *Quarterly Journal of Economics* 124(3):1095–1131.

Hero, Rodney E. 2003. "Social Capital and Racial Inequality in America." *Perspectives on Politics* 1(1):113–122.

Holland, J., C. Ramazanoglu, S. Scott, S. Sharpe and R. Thomson. 1992. "Risk, Power and the Possibility of Pleasure: Young Women and Safer Sex." *AIDS Care* 4(3):273–283.

Iqbal, Zaryab and Christopher Zorn. 2010. "Violent Conflict and the Spread of HIV/AIDS in Africa." *Journal of Politics* 72(1):149–162.

Jennifer, Kates, Adam Wexler and Eric Lief. 2016. Financing the Response to HIV in Low- and Middle-Income Countries: International Assistance from Donor Governments in 2015. The Henry J. Kaiser Family Foundation and UNAIDS Report No. 7347-12.

Jones, Kate E., Nikkita G. Patel, Marc A. Levy, Adam Storeygard, Deborah Balk, John L. Gittleman and Peter Daszak. 2008. "Global Trends in Emerging Infectious Diseases." *Nature* 451(7181):990–993.

Kalmijn, Matthijs. 1994. "Assortative Mating by Cultural and Economic Occupational Status." *American Journal of Sociology* 100(2):422–452.

Keesing, Donald B. and Donald R. Sherk. 1971. "Population Density in Patterns of Trade and Development." *American Economic Review* 61(5):956–961.

Kenyon, Chris and Robert Colebunders. 2013. "Birds of a Feather: Homophily and Sexual Network Structure in Sub-Saharan Africa." *International Journal of STD and AIDS* 24(3):1–5.

Kesler, Christel and Irene Bloemraad. 2010. "Does Immigration Erode Social Capital? The Conditional Effects of Immigration-Generated Diversity on Trust, Membership, and Participation Across 19 Countries, 1981-2000." *Canadian Journal of Political Science* 43(2):319–347.

Kimball, Ann Marie. 2006. *Risky Trade: Infectious Disease in the Era of Global Trade.* Abingdon: Routledge.

Klitsch, M. 1992. "Rural Ugandan Women's HIV Infection Rates Seem Related to Trade Routes." *International Family Planning Perspectives* 18(2):79–79.

Knack, Stephen and Philip Keefer. 1997. "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation." *Quarterly Journal of Economics* 112(4):1251–1288.

Kocsor, Ferenc, Rita Rezneki, Szabolcs Juhász and Tamás Bereczkei. 2011. "Preference for Facial Self-Resemblance and Attractiveness in Human Mate Choice." *Archives of Sexual Behavior* 40(6):1263–1270.

Kremer, Michael. 1996. "Integrating Behavioral Choice into Epidemiological Models of AIDS." *Quarterly Journal of Economics* 111(2):549–573.

Laumann, Edward O., John H. Gagnon, Robert T. Michael and Stuart Michaels. 1994. *The Social Organization of Sexuality.* Chicago: University of Chicago Press.

Laumann, Edward and Yoosik Youm. 1999. "Racial/Ethnic Group Differences in the Prevalence of Sexually Transmitted Diseases in the United States: A Network Explanation." *Sexual Transmitted Diseases* 26(5):250–261.

Le Bon, Olivier, Michel Hansenne, D. Amaru, Adelin Albert, M. Ansseau and Sam Dupont. 2013. "Assortative Mating and Personality in Human Couples: A Study Using Cloningers Temperament and Character Inventory." *Psychology* 4(1):1–8.

Lear, Dana. 1995. "Sexual Communication in the Age of AIDS: The Construction of Risk and Trust Among Young Adults." *Social Science and Medicine* 41(9):1311–1323.

Levine, Sheen S., Evan P. Apfelbaum, Mark Bernard, Valerie L. Bartelt, Edward J. Zajac and David Stark. 2014. "Ethnic Diversity Deflates Price Bubbles." *Proceedings of the National Academy of Sciences* 111(52):18524–18529.

Lin, Faqin and Nicholas C.S. Sim. 2013. "Trade, Income and the Baltic Dry Index." *European Economic Review* 59(Supplement C):1–18.

Lin, Faqin and Nicholas C.S. Sim. 2015. "Exports, HIV Incidence and the Baltic Dry Index: Further Evidence from Sub-Saharan Africa." *Economics Letters* 126(Supplement C):35–39.

Mah, Timothy L. and James D. Shelton. 2011. "Concurrency Revisited: Increasing and Compelling Epidemiological Evidence." *Journal of the International AIDS Society* 14(33):1–8.

Markel, Howard. 2003. "HIV Secrecy Is Proving Deadly." *The New York Times.* pg. F6.

Marshall, Monty G. and Keith Jaggers. 2009. *Polity IV Project: Dataset Users' Manual.* Center for Global Policy, George Mason University.

Masanjala, Winford. 2007. "The Poverty-HIV/AIDS Nexus in Africa: A Livelihood Approach." *Social Science and Medicine* 64(5):1032–1041.

Milinski, Manfred. 2006. "The Major Histocompatibility Complex, Sexual Selection, and Mate Choice." *Annual Review of Ecology, Evolution, and Systematics* 37(1):159–186.

Milner, Helen V. and Keiko Kubota. 2005. "Why the Move to Free Trade? Democracy and Trade Policy in the Developing Countries." *International Organization* 59(1):107–143.

Mitchell, Jason W. and Keith J. Horvath. 2013. "Factors Associated with Regular HIV Testing Among a Sample of US MSM with HIV-Negative Main Partners." *Journal of Acquired Immune Deficiency Syndromes* 64(4):417–423.

Montalvo, Jose G. and Marta Reynal-Querol. 2002. Why Ethnic Fractionalization? Polarization, Ethnic Conflict and Growth. UPF Economics Working Papers No. 660.

Montalvo, Jose G. and Marta Reynal-Querol. 2005. "Ethnic Diversity and Economic Development." *Journal of Development Economics* 76(2):293–323.

Montalvo, Jose G. and Marta Reynal-Querol. 2007. "Fighting Against Malaria: Prevent Wars While Waiting for the Miraculous Vaccines." *Review of Economics and Statistics* 89(1):165–177.

Morris, Chester N. and Alan G. Ferguson. 2006. "Estimation of the Sexual Transmission of HIV in Kenya and Uganda on the Trans-Africa Highway: The Continuing Role for Prevention in High Risk Groups." *Sexually Transmitted Infections* 82(5):368–371.

Morris, Martina. 1991. "A Log-Linear Modeling Framework for Selective Mixing." *Mathematical Biosciences* 107(2):349–377.

Mugweni, Esther, Stephen Pearson and Mayeh Omar. 2015. "Concurrent Sexual Partnerships Among Married Zimbabweans: Implications for HIV Prevention." *International Journal of Womens Health* 2015(7):819–832.

Newey, Whitney K. and Kenneth D. West. 1987. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55(3):703–708.

Nunn, Nathan and Nancy Qian. 2014. "US Food Aid and Civil Conflict." *American Economic Review* 104(6):1630–1666.

Nunnenkamp, Peter and Hannes Ohler. 2011. "Throwing Foreign Aid at HIV/AIDS in Developing Countries: Missing the Target?" *World Development* 39(10):1704–1723.

Oster, Emily. 2012. "Routes of Infection: Exports and HIV Incidence in Sub-Saharan Africa." *Journal of the European Economic Association* 10(5):1025–1058.

Pamuk, Şevket. 2007. "The Black Death and the Origins of the 'Great Divergence' Across Europe, 1300-1600." *European Review of Economic History* 11(3):289–317.

Pearson, Karl. 1903. "Assortative Mating in Man: A Cooperative Study." *Biometrika* 2(4):481–489.

Perrett, David I., Ian S. Penton-Voak, Anthony C. Little, Bernard P. Tiddeman, D. Michael Burt, Natalie Schmidt, Roz Oxley, Nicholas Kinloch and Louise Barrett. 2002. "Facial Attractiveness Judgements Reflect Learning of Parental Age Characteristics." *Proceedings: Biological Sciences* 269(1494):873–880.

Pesaran, Hashem M. 2004. General Diagnostic Tests for Cross Section Dependence in Panels. IZA Discussion Paper No. 1240.

Pesaran, Hashem M. 2006. "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure." *Econometrica* 74(4):967–1012.

Pitpitan, Eileen V., Seth C. Kalichman, Lisa A. Eaton, Kathleen J. Sikkema, Melissa H. Watt and Donald Skinner. 2012. "Gender-Based Violence and HIV Sexual Risk Behavior: Alcohol Use and Mental Health Problems as Mediators Among Women in Drinking Venues, Cape Town." *Social Science and Medicine* 75(8):1417–1425.

Pivnick, Anitra. 1993. "HIV Infection and the Meaning of Condoms." *Culture, Medicine and Psychiatry* 17(4):431–453.

Poulin, Michelle. 2007. "Sex, Money, and Premarital Partnerships in Southern Malawi." *Social Science and Medicine* 65(11):2383–2393.

Putnam, Robert D. 2007. "E Pluribus Unum: Diversity and Community in the Twenty-First Century." *Scandinavian Political Studies* 30(2):137–174.

Ramachandran, Sohini, Omkar Deshpande, Charles C. Roseman, Noah A. Rosenberg, Marcus W. Feldman and Luca L. Cavalli-Sforza. 2005. "Support from the Relationship of Genetic and Geographic Distance in Human Populations for a Serial Founder Effect Originating in Africa." *Proceedings of the National Academy of Sciences of the United States of America* 102(44):15942–15947.

Rushton, J. Philippe. 1988. "Genetic Similarity, Mate Choice, and Fecundity in Humans." *Ethology and Sociobiology* 9(1):329–333.

Serwadda, David, Maria J. Wawer, Stanley D. Musgrave, Nelson K. Sewankambo, Jonathan E. Kaplan and Ronald H. Gray. 1992. "HIV Risk Factors in Three Geographical Strata of Rural Rakai District, Uganda." *AIDS* 9(6):983–989.

Silventoinen, Karri, Jaakko Kaprio, Eero Lahelma, Richard J. Viken and Richard J. Rose. 2003. "Assortative Mating by Body Height and BMI: Finnish Twins and Their Spouses." *American Journal of Human Biology* 15(5):620–627.

Smith, Daniel J. 2007. "Modern Marriage, Men's Extramarital Sex, and HIV Risk in Southeastern Nigeria." *American Journal of Public Health* 97(6):997–1005.

Soroka, Stuart N., Richard Johnston and Keith Banting. 2007. Ethnicity, Social Capital and the Welfare State. In *Social Capital and theWelfare State*, ed. Fiona M. Kay and Richard Johnston. Vancouver: BC: University of British Columbia Press pp. 95–132.

Spiegel, Paul B., Anne R. Bennedsen, Johanna Claass, Laurie Bruns, Njogu Patterson, Dieudonne Yiweza and Marian Schilperoord. 2007. "Prevalence of HIV Infection in Conflict-Affected and Displaced People in Seven Sub-Saharan African Countries: A Systematic Review." *Lancet* 369(9580):2187–2195.

Spolaore, Enrico and Romain Wacziarg. 2009. "The Diffusion of Development." *Quarterly Journal of Economics* 124(2):469–529.

Spuhler, James N. 1968. "Assortative Mating with Respect to Physical Characteristics." *Eugenics Quarterly* 15(2):128–140.

Stein, Michael D., Freedberg A. Kenneth, Sullivan M. Lisa, Jacqueline Savetsky, Levenson M. Suzette, Hingson Ralph and Samet M. Jeffrey. 1998. "Sexual Ethics: Disclosure of HIV-Positive Status to Partners." *Archives of Internal Medicine* 158(3):253–257.

Steinbrook, Robert. 2007. "HIV in India - A Complex Epidemic." *New England Journal of Medicine* 356(11):1089–1093.

Syvertsen, Jennifer L., Angela Robertson Bazzi, Gustavo Martinez, Gudelia M. Rangel, Monica D. Ulibarri, Kirkpatrick B. Fergus, Hortensia Amaro and Steffanie A. Strathdee. 2015. "Love, Trust, and HIV Risk Among Female Sex Workers and Their Intimate Male Partners." *American Journal of Public Health* 105(8):1667–1674.

Tavory, Iddo and Ann Swidler. 2009. "Condom Semiotics: Meaning and Condom Use in Rural Malawi." *American Sociological Review* 74(2):171–189.

UNAIDS. 2009. The Global Economic Crisis and HIV Prevention and Treatment Programmes: Vulnerabilities and Impact. UNAIDS and World Bank Report.

UNAIDS. 2014. *The Gap Report*. Geneva: Joint United Nations Programme on HIV/AIDS.

UNAIDS. 2017. *Ending AIDS: Progress Towards the 90-90-90 Targets*. Geneva: Joint United Nations Programme on HIV/AIDS.

UNHCR. 2007. Refugees and HIV. UNHCR Plus News.

USAID. 2003. HIV/AIDS and Conflict. USAID Fact Sheet.

Zietsch, Brendan P., Karin J. H. Verweij, Andrew C. Heath and Nicholas G. Martin. 2011. "Variation in Human Mate Choice: Simultaneously Investigating Heritability, Parental Influence, Sexual Imprinting, and Assortative Mating." *American Naturalist* 177(5):605–616.

Appendix A: Supplementary Results

Oster (2012) finds that in sub-Saharan Africa, an increase in exports can lead to an increase in HIV incidence. In this appendix, we provide additional regression results that are based on using the log of exports per capita (instead of trade) as the treatment in our difference-in-differences approach.

In Table A1, we re-estimate our model with the log of exports per capita, and also with the control variables (one at the time or all at once) used in the sensitivity analysis presented in Section 3.5.3. As before, we find that the increase in exports may lead to an increase in HIV incidence in the SSAs. Moreover, in all the columns of Table A1, the interaction between the log of exports per capita and genetic diversity has a negative and statistically significant coefficient. This suggests that genetic diversity can mitigate the positive effect of exports on HIV incidence, and that this mitigating effect is present whether we use trade or exports as the treatment.

We also repeat the mechanism analysis presented in Section 3.5.4 with log of exports per capita as the treatment variable. Like our baseline regressions, Table A2 shows that genetic diversity, cultural diversity, and (the lack of) trust would reduce the impact of exports on HIV incidence. Therefore, genetic homophily, cultural homophily and trust are once again possible mechanisms underlying the negative impact of genetic diversity on HIV.

Table A1: Exports, Genetic Diversity and HIV Incidence: Robustness Estimates

	Dependent Variable: HIV Incidence (%)										
Method: CCE	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Log exports per capita	0.376***	0.286***	0.416***	0.548***	0.320**	0.353***	0.541***	0.375***	0.335***	0.375***	0.713***
	(0.098)	(0.085)	(0.125)	(0.156)	(0.147)	(0.101)	(0.096)	(0.104)	(0.097)	(0.093)	(0.241)
Log exports per capita×	−0.441***	−0.356***	−0.493***	−0.669***	−0.384**	−0.405***	−0.667***	−0.441***	−0.379***	−0.423***	−0.818***
genetic diversity	(0.126)	(0.112)	(0.161)	(0.202)	(0.187)	(0.130)	(0.123)	(0.134)	(0.125)	(0.113)	(0.300)
Log GDP per capita$_{t-1}$		0.044***									−0.062***
		(0.009)									(0.009)
Log ODA per capita$_{t-1}$			0.013								0.046***
			(0.011)								(0.009)
Democracy$_{t-1}$				0.007***							−0.001
				(0.002)							(0.002)
Primary enrollment$_{t-1}$					0.002***						0.003***
					(0.001)						(0.001)
Log population density$_{t-1}$						0.015***					0.051***
						(0.003)					(0.009)
Conflict$_{t-1}$							−0.083***				−0.075***
							(0.009)				(0.006)
Malaria intensity$_{t-1}$								0.000			−0.003
								(0.003)			(0.003)
Log rainfall$_{t-1}$									−0.072***		−0.170***
									(0.012)		(0.050)
Commodity price$_{t-1}$										−0.216**	−0.072
										(0.099)	(0.135)
HIV prevalence$_{t-1}$	0.119***	0.119***	0.119***	0.118***	0.118***	0.119***	0.118***	0.119***	0.117***	0.117***	0.110***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
Time trend (general)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Interactive fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	927	927	918	927	773	927	927	927	927	927	766
Number of countries	36	36	36	36	36	36	36	36	36	36	36

Note: Newey-West standard errors are reported in the parenthesis. The data on HIV incidence are available in 37 SSAs, while 36 countries of which have the data on genetic diversity. Statistical significance at the 10%, 5% and 1% levels are indicated by *, **, and ***, respectively.

Table A2: Genetic Diversity, Cultural Diversity, Trust and HIV Incidence

	Dependent Variable: HIV Incidence (%)			
Method: CCE	(1)	(2)	(3)	(4)
Trust measure:	Overall	Relatives	Neighbors	Others
Log exports per capita	1.984***	1.614**	1.874***	1.877***
	(0.577)	(0.692)	(0.549)	(0.484)
Log exports per capita × genetic diversity	−2.341***	−1.911**	−2.214***	−2.206***
	(0.802)	(0.933)	(0.729)	(0.644)
Log exports per capita × religious diversity	−0.122	−0.118	−0.112*	−0.116*
	(0.079)	(0.076)	(0.064)	(0.067)
Log exports per capita × ethnic diversity	0.005	−0.012	0.028	0.034
	(0.011)	(0.024)	(0.023)	(0.022)
Log exports per capita × linguistic diversity	0.164***	0.159***	0.161***	0.160***
	(0.027)	(0.026)	(0.033)	(0.037)
Log exports per capita × trust	0.053***	0.024**	0.007	0.004
	(0.014)	(0.011)	(0.013)	(0.016)
HIV prevalence$_{t-1}$	0.092***	0.091***	0.092***	0.092***
	(0.005)	(0.005)	(0.005)	(0.005)
All controls	Yes	Yes	Yes	Yes
Time trend (general)	Yes	Yes	Yes	Yes
Interactive fixed effects	Yes	Yes	Yes	Yes
Number of observations	508	508	508	508
Number of countries	23	23	23	23

Note: Newey-West standard errors are reported in the parenthesis. All regressions control for income, foreign aid, democracy, primary enrollment rate, population density, armed conflict, malaria intensity, rainfall, and commodity price. The religious, ethnic and linguistic diversity data are available for 31, 30 and 35 SSA countries, respectively. Although the data on trust taken from the Afrobarometer survey of Round 5 are available for 35 African countries, only 26 of which have data on HIV incidence. Overall, 23 SSAs have the data on genetic diversity, religious, ethnic and linguistic diversity and trust. Statistical significance at the 10%, 5% and 1% levels are indicated by *, **, and ***, respectively.

Appendix B: A Model of Population Diversity and the HIV/AIDS

B.1 Set-up

In epidemiology, models of HIV/AIDS often depict new cases of HIV infections as positively associated with the number of sex partners (see, for example, Kremer 1996; Garnett 2002). As such, conditional on the same sexual practice, a person with a larger number of sex partners in a given period would be at higher risk of contracting HIV than a person with fewer sex partners. In this paper, we hypothesize that the diversity in physical traits as a result of genetic differences and cultural heterogeneity are barriers to forming sexual relationships. In other words, the cost of forming sexual relationships would be higher among individuals who are more genetically and culturally dissimilar, and therefore, in societies with lower levels of trust or those that exhibit intra-group sexual preference.

For our conceptual model, we consider a country with a sexually active population of n, of which α % is HIV-positive (i.e. α is the HIV prevalence rate) and δ % is susceptible. We assume that the probability of HIV transmission per sexual partnership is β. In principle, β could be influenced by the type of sexual practice (e.g. protected or unprotected sex), but to keep things simple, let us fix β to be the transmission rate.[36] In addition, let s_i be a number of sex partners person i ($i = 1, 2, ..., n$) has had for a given period of time, which is normalized to one for simplicity. The number of sex partners one has is related to the average cost of having a sex partner, w_i and income m_i. For instance, suppose that the person i has l number of sex partners, and the cost to person i when forming a sexual partnership with person k is w_{ik}, $k = 1, 2, ..., l$. Hence, we may write the average cost of having a sex partner as $w_i = \left[\sum_{k=1}^{l} w_{ik}\right]/l$. Poulin (2007) (as cited in Greenwood et al. 2013) provides evidence of expenditure on partners in the form of gifts in exchange for sex in Malawi. This may apply to other African countries as well when sex comes with the transfers mostly from men to women. On the aggregate level, each individual in the country therefore on average has a number of sex partners given by $\bar{s} = \frac{1}{n}\sum_{i=1}^{n} s_i(w_i, m_i)$, the average cost of sex partners, $\bar{w} = \sum_{i=1}^{n} w_i$ and the average income, $\bar{m}$. Following epidemiological model of sexually-transmitted disease epidemic

[36] We can write $\beta_{unprotect} > \beta_{protect}$.

(see e.g., Kremer 1996; Garnett 2002), the new infection rate ρ now is[37]

$$\rho = f\big(\alpha, \delta, \beta, \bar{s}(\bar{w}, \bar{m})\big), \tag{B.1}$$

which is increasing in α, δ, β and $\bar{s}$. Initially, if the prevalence rate, α is high the chance of the susceptible group, δ having sexual relationship with someone infected is also high, hence increasing new cases of HIV infection. Likewise, having sex with an infected partner does not always cause one to get infected, yet the chance of transmission, β varies depending on various factors including sexual practices of the person. For instance, homosexual sex incurs a person a very high risk compared to heterosexual sex does (CDC 2016). Finally, having multiple sex partners, all other factors being fixed, also puts individuals at a higher risk of being infected, leading to an increase in incidence.

What makes equation (1) different from the epidemiological model is that we allow $\bar{s}$ to depend on $\bar{w}$ and $\bar{m}$. In other words, the cost of having sex partners as well as income influence the behavior of individuals to have many or few partners. If α, δ and β are assumed to be exogenous, we are interested in the effect of the change in $\bar{w}$ and $\bar{m}$ on ρ. Mathematically, the effect is

$$\frac{\partial \rho}{\partial \bar{w}} = \frac{\partial f(.)}{\partial \bar{s}} \frac{\partial \bar{s}}{\partial \bar{w}} \tag{B.2}$$

and

$$\frac{\partial \rho}{\partial \bar{w}} = \frac{\partial f(.)}{\partial \bar{s}} \frac{\partial \bar{s}}{\partial \bar{m}} \tag{B.3}$$

Provided that the first terms on the right-hand size of equation B.2 and B.3 are always positive, the effect of the cost or income on the incidence follows the sign of $\frac{\partial \bar{s}}{\partial \bar{w}}$ or $\frac{\partial \bar{s}}{\partial \bar{m}}$, respectively. To analyze the relationship between $\bar{s}$ with $\bar{w}$ and $\bar{m}$, we use the following setup.

In the same country, an individual i's preference is presented by a utility which is additive separable between number of sex partners s_i and consumption c_i. We assume that $U(s_i)$ and $U(c_i)$ are strictly concave, which implies that $U'_{s_i}(.) > 0$, $U''_{s_i}(.) < 0$, $U'_{c_i}(.) > 0$, and $U''_{c_i}(.) < 0$, so here we rule out the case of abstinence where $U'_{s_i}(.) < 0$. Moreover, by engaging in sexual activity, the individual already knows that there is still a chance of contracting HIV. Let us denote the perceived probability of infection

[37] For simplicity, we ignore the HIV/AIDS related death rate in the model.

by π_i. Based on Duflo, Dupas and Kremer (2015)[38], we simplify π_i to be a function of number of sex partners, s_i and that $\pi'_{s_i} > 0$ and $\pi''_{s_i} < 0$. That means the individual does perceive that he or she would have a high chance of getting infected if having multiple partners. Contracting HIV then costs the individual a loss in utility given by F_i. The individual i is endowed with income m_i which is exhausted through consumption, c_i and expenditure on having sex partners, $s_i w_i$. A familiar budget constraint is given by $c_i + w_i s_i = m_i$. The total utility of person i with consumption c_i and number of sex partners s_i is

$$V_i(s_i, c_i) = U(s_i) + U(c_i) - \pi_i(s_i)F_i \tag{B.4}$$

The optimal s_i^* and c_i^* are derived by solving the following first-order necessary condition

$$\frac{\partial U_i(s_i)}{\partial s_i} - \frac{\partial \pi_i(s_i)}{\partial s_i}F_i = \frac{\partial U_i(c_i)}{\partial c_i}w_i \tag{B.5}$$

Equation B.5 shows the trade-off between consumption and number of sexual partners. On the left-hand size, if the person i increases one more partner, he or she enjoys the increasing utility given by $\partial U_i(.)/\partial s_i$, but faces the risk of contracting HIV with the expected loss in utility by $(\partial \pi_i/\partial s_i)F_i$. The term on the right-hand size, $(\partial U_i(.)/\partial c_i)w_i$ is the loss in utility due to the decrease in consumption since money has been spent on the extra sexual partner.

To ensure the maximum, the relevant border Hessian determinant must be positive. We can derive the bordered Hessian matrix as

$$H = \begin{bmatrix} 0 & w_i & 1 \\ w_i & f_{ss} & 0 \\ 1 & 0 & f_{cc} \end{bmatrix},$$

with f_{ss} and f_{cc} being the second direct partial derivatives and using the fact that the second cross partial derivatives, f_{sc} and f_{cs} are equal to 0. The determinant of H is computed to be $-f_{ss} - w_i^2 f_{cc}$. Therefore, we must have

$$\frac{\partial^2 U(c_i)}{\partial c_i^2}w_i^2 - \left[\frac{\partial^2 U(s_i)}{\partial s_i^2} - \frac{\partial^2 \pi_i(s_i)}{\partial s_i^2}F_i\right] > 0 \tag{B.6}$$

[38]They allow the perceived probability of infection to depend on the number of sexual acts and the perceived chance of infection per unprotected sex act.

B.2 A Link to Sexual Preference and Trust

First, suppose that a distance between a pair of individuals in the population with respect to for instance physical appearance or culture can be represented in a real line and is linearly ordered. Let $d_{ik} = |i - k|$ $(i, k \in \mathbb{R}^+)$ denote the absolute distance between individual i and his/her partner k, so the distance between individual number 1 and 3 is closer to that between individual number 10 and 15 and so on. We further assume that matching with a partner who is greatly dissimilar incurs the individual higher cost than matching with one who is relatively similar. One example is when a person tries to start a relationship with a partner who comes from a different ethnic group and speaks a different language. In such a case, the person may have to spend more resources including a financial resource to build the relationship in order to get into sex. This reason however may be less convincing if the individual chooses to have sex with prostitutes since he or she just settles one-off payment and enjoys sex. However, there is also anecdotal evidence that foreigners have to pay more than the locals when they want to have sex with local prostitutes. This is very much due to the cultural distance between them.

Moreover, as we argued previously, in a highly diverse country, the average distance among individuals in the population is greater than that in a less heterogeneous one. In this sense, great distance reduces sexual preference and lowers trust. Building on these assumptions and again supposing that the individual i has l number of sex partners, we can impose a linear relationship between the distance and average cost of having sex partners as follows

$$w_i = \frac{\theta_i \sum_{k=1}^{l} d_{ik}}{l} \tag{B.7}$$

where the constant $\theta_i > 0$, so $\frac{\partial w_i}{\partial d_{ik}} > 0$. Therefore, in this set-up trust is just the cost of having a sex partner, which is assumed to vary depending on the physical and cultural distance.

B.3 Comparative Statics

Next, we use the comparative static to see the effect of the change in cost and income on the change in number of sex partners and subsequently the HIV incidence.

Proposition 1: *If the average cost of having sex partners increases, individuals tend to have fewer partners. Basically,* $\frac{\partial s_i}{\partial w_i} < 0$.

Proof. Differentiate the first order condition in B.5 by w_i to get

$$\frac{\partial s_i}{\partial w_i} = \frac{\dfrac{\partial^2 U(c_i)}{\partial c_i^2}\dfrac{\partial c_i}{\partial w_i}w_i + \dfrac{\partial U(c_i)}{\partial c_i}}{\dfrac{\partial^2 U(s_i)}{\partial s_i^2} - \dfrac{\partial^2 \pi_i(s_i)}{\partial s_i^2}F_i}$$

The nominator is positive and the denominator is negative by the property in B.6, so $\frac{\partial U(s_i)}{\partial w_i} < 0$.

Proposition 2: *Increasing income encourages individuals to have more partners if the change in income does not affect the average cost of having sex partners, w_i. Indeed, $\frac{\partial s_i}{\partial m_i} > 0$. The effect is otherwise ambiguous if the change in income causes the change in cost.*

Proof. Again, differentiate the first order condition in B.5 with respect to m_i to get

$$\frac{\partial s_i}{\partial m_i} = \frac{\dfrac{\partial^2 U(c_i)}{\partial c_i^2}\dfrac{\partial c_i}{\partial m_i}w_i + \dfrac{\partial U}{\partial c_i}\dfrac{\partial w_i}{\partial m_i}}{\dfrac{\partial^2 U(s_i)}{\partial s_i^2} - \dfrac{\partial^2 \pi_i(s_i)}{\partial s_i^2}F_i}$$

The denominator is negative by the property in B.6 and the nominator is negative if $\frac{\partial w_i}{\partial m_i} = 0$.

Proposition 3: *High/low sexual preference/trust induces individuals to have more/fewer partners, hence increasing/reducing new cases of HIV infection.*

Proof. This follows immediately using the definition of trust in B.7, Proposition 1, and the incidence in B.1.

Chapter 4

Does Geography Matter for the Spread of HIV/AIDS? Evidence from Sub-Saharan Africa

Pide Lun
School of Economics, University of Adelaide

Abstract: The roots of HIV epidemic in sub-Saharan Africa are complex. We show that the transmission of HIV in this region can be explained by a geographical factor associated with a unique historical event taking place thousands of years before. We demonstrate that countries with geographical difficulty measured by high levels of terrain ruggedness tend to have high rates of HIV infections. The positive relationship between terrain ruggedness and HIV can be established through the following conduits. Firstly, highly rugged terrain has a negative effect on the slave trade as steep or rocky topography provides protection for people being attacked and enslaved. Secondly, lower levels of the slave trade in the past leads to societies that are more trusting in the present. As trust may foster the spread of HIV, countries with highly rugged terrains may experience a more serious HIV/AIDS problem.

Keywords: HIV, sub-Saharan Africa, terrain ruggedness, slave trade, trust.

JEL Classification: I12, O10, O55.

Statement of Authorship

Title of Paper	Does Geography Matter for the Spread of HIV/AIDS? Evidence from Sub-Saharan Africa.
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted work written in manuscript style
Publication Details	This paper is written in publication style for submission to a journal.

Principal Author

Name of Principal Author (Candidate)	Pide Lun		
Contribution to the Paper	This paper is written by a sole author (the Candidate)		
Overall percentage (%)	100%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	19 January 2018

4.1 Introduction

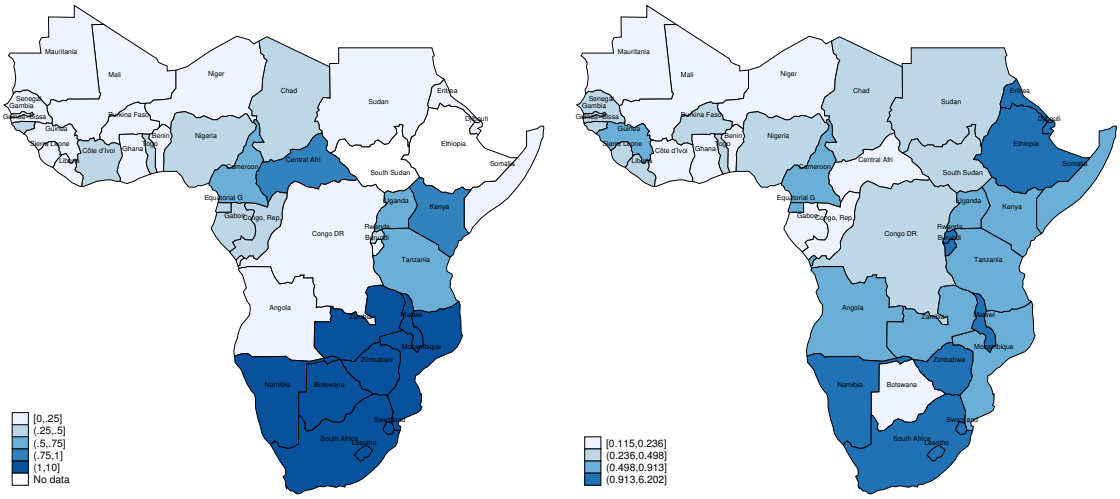
It is well known that countries in sub-Saharan Africa (*SSAs* in short) are unevenly affected by the HIV/AIDS epidemic. For instance, many SSAs in southern and eastern parts of Africa have the world's highest HIV infection rates, while other SSAs have infection rates below the world's average. Consider, for example, the HIV incidence rate, which is the new HIV infection rate among adult population. Senegal, a country in western Africa, has an HIV incidence rate of only 0.01%. By contrast, the HIV incidence rate in Zimbabwe, a country in the south, is 90 times greater. Senegal and Zimbabwe, however, are comparable in terms of economy size and population, which raises the question of whether there are deep, fundamental factors that can explain why certain SSAs are more severely affected by HIV/AIDS than others.

In this paper, we offer new evidence that the cross-sectional variation in the HIV incidence rates in sub-Saharan Africa can be linked to geography, particularly the *terrain ruggedness* of the SSAs. Terrain ruggedness is a topographical measure that characterizes how uneven an area is. Thus, countries with highly rugged terrains typically have land topography that is steeply sloping or rocky. In sub-Saharan Africa, the spatial distribution of terrain ruggedness appears to be correlated with the severity of the HIV problem. For example, using the index of terrain ruggedness constructed by Nunn and Puga (2012),[1] Figure 4.1 shows that countries with highly rugged terrains like South Africa or Zimbabwe in the south are also those with high HIV incidence rates. By contrast, countries in western Africa, where HIV/AIDS is less severe, tend to be geographically flatter.

This paper contends that there are historical reasons for why terrain ruggedness may affect the transmission of HIV, and thus, the cross-sectional variation in the HIV/AIDS epidemic across the SSAs. Our hypothesis is rooted in the seminal paper of Nunn and Puga (2012) who show that the extent of the slave trade in Africa from 1400 and 1900 can be explained by terrain ruggedness. During this period, more than 18 million slaves were abducted, mostly by force, and shipped out of Africa to the

[1]The Nunn and Puga (2012) terrain ruggedness index measures in 100 meters the average elevation score of a country's territory not covered by water, and weighted by sea-level surface.

Figure 4.1: HIV Incidence and Terrain Ruggedness



(A) HIV Incidence (%)

(B) Terrain Ruggedness Index

rest of the world (Nunn 2008).[2] Nunn and Puga (2012) argue that terrain ruggedness has a negative effect on slave export because rough topography like cliffs or caves provides protection from raids and subsequent enslavement. Areas that are highly rugged also discouraged slave capturing, as rugged terrains made it challenging for slaves to be transported to the nearest coast. Data on slave export compiled by Nunn and Puga (2012) also supports this finding by showing that during the trans-Atlantic slave trade from the 15th century, most of the slaves were taken from western Africa, such as from Nigeria and Ghana, where the geography is relatively flat. However, southern Africa, where the terrain is highly rugged, exported much fewer slaves.

Why, then, would the slave trade be relevant for the transmission of HIV in the SSAs? We argue that the reason is trust. Firstly, there is evidence that the slave trade has a negative effect on trust. Nunn and Wantchekon (2011) find that a culture of "mistrust" in Africa emerged during the period of the slave trade, particularly when slave capturing became intensified. At that time, people in the community not only had to stay vigilant of violent attacks from outsiders, but also had to be wary of any

[2]According to Nunn (2008), there are four waves of the slave trade. The trans-Atlantic slave trade was the largest one during which around 12 million slaves were shipped to America. The other three slave trades were trans-Saharan, Red Sea and Indian Ocean. During these slave trades, around 6 million slaves were taken to Northern Africa, Middle East and India. The slave trade in Africa was a tragic event because not only there was a huge number of slaves being captured and exported, but a lot of people died in bondage from brutality (Nunn 2008).

deceiving act of their close acquaintances such as friends and relatives who might want to sell them into slavery. Such neighborhood they lived in created an environment of fear and mistrust, which still persists today (Nunn and Wantchekon 2011).

When it comes to the spread of HIV, trust matters. In the epidemiology literature, trust is found to be associated with risky sexual behavior, and makes people underestimate the potential risks from sexual relationships, including the risk of contracting sexually transmitted diseases like HIV. This is evident by the fact that people are likely to forgo the use of protection, and thus expose themselves to greater risk of contracting HIV, if they are more trusting of their partners (Appleby, Miller and Rothspan 1999; Tavory and Swidler 2009; Syvertsen et al. 2015). Therefore, if terrain ruggedness has led to lower levels of the slave trade in the past, leading to societies that are more trusting in the present, the HIV/AIDS problem may turn out to be more severe in highly rugged countries.

Empirically, we test if terrain ruggedness has a positive effect on HIV incidence through the within-between estimator proposed by Allison (2009) (see also, Greene 2012; Dieleman and Templin 2014; Bell and Jones 2015). This approach is necessitated by the fact that terrain ruggedness is time-invariant; thus, we cannot observe its impact on HIV incidence once country fixed effects are used. The within-between estimator, also known as hybrid approach, is a procedure that exploits both within and between variations of the regressors. In this regard, it "marries" the fixed and random effects estimators, since the former estimates the model parameters using the variables' within variation while the latter exploits their between variation. Consequently, the within-between estimator enjoys the benefits of both fixed and random effects estimators. Firstly, it produces within-effect estimates that are identical to the estimates produced by the fixed effects estimator. Secondly, like the random effects estimator, it is able to estimate the impact of time-invariant factors, which fixed effects regressions cannot achieve (Bell and Jones 2015).

Our results show that countries with more highly rugged terrains are likely to have higher rates of HIV incidence. Our baseline result shows that on average, a one-standard deviation increase in the terrain ruggedness index (which is about 120 meters) increases HIV incidence by about 0.43 percentage point (which is more than

50% of the SSAs' average incidence rate) on average. This positive association is present even when controlling for other geographical influences such as soil quality, diamond production, tropical climate, and distance to coast, and for variables related to wars and conflicts, quality of institution and physical infrastructure, which may correlate with terrain ruggedness (Nunn and Puga 2012) and HIV incidence. Moreover, it is also robust to the use of alternative HIV incidence data reconstructed based on the method of Oster (2012), alternative geographical measures of terrain ruggedness, and to the use of Bayesian Model Averaging to account for potential uncertainty in model specification.

We show that the historical slave trade is one mechanism through which terrain ruggedness affects HIV incidence. Specifically, we introduce a variable on slave export into the baseline regression. We find that slave export has a negative effect on HIV incidence, and more importantly, it *drives out* the statistically significant effect of terrain ruggedness previously observed. That the effect of terrain ruggedness on HIV incidence can be explained by slave export is consistent with the notion that societies with higher levels of the historical slave trade have lower levels of trust today, and the lack of trust, in turn, constricts the spread of HIV.

This paper is related to three themes intersecting economic geography, history, and health. Firstly, it is related to studies that look at the impact of geography on economic outcomes. Disadvantageous geography is argued to be an obstacle to economic progress (Gallup, Sachs and Mellinger 1999; Sachs, Mellinger and Gallup 2001; Rappaport and Sachs 2003). Along these lines, we provide new evidence that poor geography like rugged terrains may also affect development in sub-Saharan Africa negatively through its impact on the transmission of HIV/AIDS. Secondly, this paper is related to the literature that speaks to the role of history in contemporary economic development (Acemoglu, Johnson and Robinson 2001; Banerjee and Iyer 2005; Nunn and Wantchekon 2011; Alesina, Giuliano and Nunn 2013). We show that the interaction of geography with history (i.e. the slave trade) may affect people's risky sexual behavior, which may then explain why rugged SSAs have higher HIV incidence rates. Finally, this study is part of the broad literature that explores HIV risk factors and determinants, especially in the context of sub-Saharan Africa (Kremer

1996; Lakdawalla, Sood and Goldman 2006; Smith 2007; Swidler and Watkins 2007; de Walque 2007, 2009; UNAIDS 2010; Oster 2012; Duflo, Dupas and Kremer 2015; Lin and Sim 2015).

The remainder of the paper is organized as follows. Next section describes the data and summary statistics of key variables used in the analysis. We discuss the proposed estimation strategy in Section 4.3, followed by the empirical results and discussion in Section 4.4. Section 4.5 concludes.

4.2 Data Sources

Our work uses two main datasets: longitudinal data on HIV incidence rates collected by UNAIDS, and data on terrain ruggedness and the slave trade compiled by Nunn and Puga (2012). Data on control variables employed in our analysis are obtained from multiple sources.

HIV Incidence The outcome variable of interest is *HIV incidence*, defined as the percentage of new HIV infections among adult population aged between 15 and 49 in a given year. We retrieve the data on HIV incidence rates for the SSAs from UNAIDS for the period between 1990 and 2014.[3] The country-level HIV incidence rates published by UNAIDS are modeled estimates based on HIV data obtained from antenatal clinics, population-based surveys, and repeated HIV prevalence studies (UNAIDS 2016). For our robustness checks, we also consider the HIV incidence data reconstructed using the method proposed by Oster (2012). This method allows one to estimate the HIV incidence rates for the years before 1990, during which UNAIDS data are not available. The HIV incidence rate at time t computed by Oster is defined as $I_t = P_t - P_{t-1} + \sum_{i=1}^{t-1} d_{t-i} I_i$, where P_t is the prevalence in year t and d_{t-i} is the likelihood of dying from AIDS after $t-i$ years of infection. To compute the prevalence rates since the beginning of the epidemic, Oster uses data on HIV infections recorded in antenatal clinics published in UNAIDS country reports and employs a linear interpolation method. The approximate time to death from AIDS

[3] See http://aidsinfo.unaids.org for the data source.

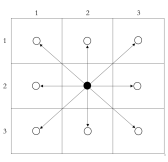
is estimated using data from developed countries, yet it is adjusted to account for the fact that the time to death is faster in developing countries (i.e. in Africa) than in developed countries. Using Oster method, we may construct a dataset on HIV incidence rates for the SSAs from 1985 to 2014.

Terrain Ruggedness Our data on *terrain ruggedness* is taken from Nunn and Puga (2012). They, in turn, compute the index of terrain ruggedness using a method by Riley, Degloria and Elliot (1999). For a given location, this index measures the aggregated variation in elevation between each point within this location and its eight neighboring points in compass directions.[4] Specifically, it is calculated as $\sqrt{\sum_{i=r-1}^{r+1}\sum_{i=c-1}^{c+1}(e_{i,j}-e_{r,c})^2}$, where $e_{r,c}$ is the elevation score of a point in row r and column c of a grid.[5] The data of elevation scores are extracted from the US Geological Survey. Nunn and Puga then compute a country-level terrain ruggedness index, which is the average index of all location points measured in 100 meters within each country's territory not covered by water, and weighted by the sea-level surface of the earth's oceans.

In our sensitivity analysis, we use four alternative measures of terrain ruggedness, suggested by Nunn and Puga (2012), to check if our results are sensitive to the way ruggedness is measured. The first is the *population-weighted terrain ruggedness index*, which is the terrain ruggedness index (i.e. *terrain ruggedness*) adjusted for population density. The second is the *average absolute slope*, which is the mean of the absolute differences in the slope between a location point and the eight neighboring points in compass directions. The third is the *average standard deviation of elevation* between a location point and the eight surrounding points. The last is the *rate of highly rugged terrain*, defined as the proportion of country's land area covered by rugged

[4]Those are north, northeast, east, southeast, south, southwest, west, and northwest.

[5]For example, in a cell grid (adapted from Nunn and Puga (2012)) below, the point in the northwest of the central location is in row 1 of column 1. The distance between location points in the grid is 0.5 nautical mile, equaling to 926 meters.



terrain measured above 250 meters. If rough geography affects HIV infections, our conclusion that terrain ruggedness affects HIV incidence should not be affected by how ruggedness is measured.

Slave Trade Our data on the slave trade is taken from Nunn and Puga (2012). Specifically, information on the slave trade is captured by slave export, which is defined as the total number of slaves exported from the SSAs to the rest of the world between 1400 and 1900 during the four waves of the slave trades, namely the trans-Atlantic, Indian Ocean, Red Sea, and trans-Saharan. The original slave export data was first compiled from several sources and published by Nunn (2008). During the course of this slave trade, countries such as Angola, Nigeria, Ghana and Ethiopia exported the highest number of slaves of over 8 million, whereas Lesotho, Swaziland or Botswana did not ship a single slave (Nunn 2008). In the regression analysis, we use *slave export intensity* defined as $\log\left(1+\frac{slave\,export\times 100}{land\,area}\right)$ which adjusts for country size, and accounts for the fact that the distribution of slave export is highly skewed and that some countries did not ship any slaves as mentioned (Nunn 2008).

Controls Our control variables and their sources are as follows. *GDP per capita* (in log) measured in 2005 US dollars and *population density* (in log) are extracted from the World Bank's World Development Indicators. For geographical factors, we follow Nunn and Puga (2012) in employing four variables: *soil quality*, defined as the proportion of land area with fertile soil (which is soil that is suitable for all rain-fed crops); *diamonds*, defined as the total carats of gem-quality diamond extracted from 1958 to 2000 normalized by country size; *tropical climate*, which is the percentage of land surface area with any of Köppen-Geiger tropical climates;[6] and *distance to coast*, defined as the average distance (measured in thousand kilometers) to the nearest ice-free coast for each country.

Aside from these controls, we consider three control variables related to conflict, quality of institution and infrastructure, all of which may confound the effect of terrain ruggedness on HIV incidence. For conflict, we exploit three indicator

[6]The climate classifications are rainforest, monsoon and savanna (wet and dry).

Table 4.1: Descriptive Statistics of Main Variables

	Mean (Standard Deviation)			
	South	East	West	Central
HIV incidence (%)	1.934	0.489	0.194	0.467
	(1.228)	(0.589)	(0.169)	(0.324)
Terrain ruggedness (100 m.)	1.702	1.438	0.281	0.328
	(1.786)	(0.946)	(0.164)	(0.173)
Terrain ruggedness (population-weighted, 100 m.)	1.272	1.092	0.318	0.640
	(1.177)	(0.848)	(0.235)	(0.401)
Average absolute slope (%)	4.741	3.822	0.745	0.868
	(5.089)	(2.506)	(0.434)	(0.450)
Standard deviation of elevation (100 m.)	0.435	0.387	0.082	0.103
	(0.446)	(0.241)	(0.047)	(0.054)
Rate of highly rugged terrain (%)	22.716	20.820	0.706	2.056
	(28.698)	(19.543)	(1.064)	(2.292)
Slave export (per 10,000 squared m.)	4.146	3.845	19.069	1.061
	(9.124)	(5.060)	(21.682)	(1.028)

Note: Number of observations: South = 225; East = 250; West = 325; Central= 125.

variables that include *any war* (i.e. if a country had experienced a conflict[7]), *intra-state conflict* (i.e. if a country had experienced a conflict between a government and a non-governmental group without any involvement of other nations) and *inter-state conflict* (i.e. if a country had experienced a conflict between two or more governments), and a category variable namely *conflict intensity* classified by 1 if the number of deaths is less than 1,000 and 2 if it is over 1,000. Conflict data is extracted from UCDP/PRIO[8] Armed Conflict Dataset published by the University of Uppsala. We extract from the Word Bank's World Governance Indicators, data on quality of institution measured by four indices: *rule of law*, *government effectiveness*, *voice and accountability* and *regulatory quality*. Each of these index ranges from -2.5 (lowest quality) to 2.5 (highest quality) in the scale.

To proxy for infrastructure quality, we consider *paved roads*, defined as the proportion of paved roads to total roads, which can be extracted from the World bank's World Development Indicators. We use average proportion of paved roads between 1990 and 2000 in the regressions because annual data on paved roads published by the World Bank contains significant missing observations across the years for the SSAs.

[7]It is categorized as a conflict if the number of battle-related deaths is at least 25.

[8]Uppsala Conflict Data Program/Peace Research Institute.

Table 4.1 presents the summary statistics of the main variables by regions in the SSAs. It affirms that HIV incidence is high among countries in the south like Botswana and South Africa, moderate in the east and central, and low in the west. The uneven distribution of HIV incidence is correlated with the distribution of rough geography, where more rugged terrains can be found in the south and east than in the west and central. However, the distributions of both HIV incidence and rough geography appears to be negatively correlated with the distribution of slave export. In sub-Saharan Africa, the west exported a larger number of slaves per 10,000 squared meters than the south and the east did.

4.3 Empirical Strategy

The main focus of this paper is to examine the effect of physical geography, measured by terrain ruggedness, on the spread of HIV in the SSAs. We consider estimating the following panel specification,

$$H_{it} = \alpha + \beta R_i + \mathbf{Z}'_{it}\boldsymbol{\theta} + \mathbf{X}'_i\boldsymbol{\gamma} + u_{it}, \tag{4.1}$$

where H_{it} is the HIV incidence rate of country i in year t, and R_i is the measure of terrain ruggedness. The vector $\mathbf{Z_{it}}$ represents a set of time-variant controls; these include, for instance, GDP per capita or population density. The vector $\mathbf{X}_i$ represents a set of time-invariant controls. The unobserved component u_{it} is defined as $u_{it} = a_i + b_t + \varepsilon_{it}$, where a_i and b_t represent country and year fixed effects respectively, and ε_{it} is the unobserved error term assumed to be normally distributed with zero mean and some constant variance. The coefficient of interest is β, which captures the impact of terrain ruggedness on the incidence of HIV on average. It can be interpreted as the percentage point change in HIV incidence for every 100-meter increase in terrain ruggedness.

One obvious problem when estimating the effect of terrain ruggedness on HIV incidence, based on Eq. (4.1), is the infeasibility of using country fixed effects to control for unobserved permanent differences across countries. This is because terrain

ruggedness is itself time-invariant. Given that country fixed effects will purge all time-invariant factors from the model, the effect of terrain ruggedness cannot be identified if country fixed effects are employed.

As an alternative to the fixed effects estimator, we may consider the random effects estimator, as it allows us to estimate the effects of time-invariant factors in the model. However, the random effects estimator could be biased and inconsistent if the unobserved time-invariant error component a_i is correlated with the regressors in the model. To address this problem, one could employ the Hausman-Taylor procedure (Hausman and Taylor 1981), which decomposes time-invariant and time-variant factors into exogenous and endogenous covariates. This "transformed model" is then estimated by two-stage least squares regression, assuming that there are appropriate instruments for the endogenous covariates of the model.[9] While the Hausman-Taylor approach addresses the issue of endogeneity arising from correlated fixed effects that the random effects model cannot, valid instruments are unfortunately hard to come by (Plümper and Troeger 2007).

As a compromise between the fixed and random effects estimators, this paper utilizes a method known as the *within-between estimator*, proposed by Allison (2009), to estimate the effect of time-constant factors. The method is also discussed in detail by, among others, Greene (2012), Dieleman and Templin (2014), Bell and Jones (2015). The within-between estimator is a *modified* random effects estimator, and is described as such because of its ability to estimate time-constant variables just as the random effects estimator can. However, unlike the "pure" random effects estimator, the within-between estimator corrects for endogeneity arising from the correlation of the random effects (i.e. a_i) with the time-variant factors of the model (i.e. $\mathbf{Z}_{it}$) just as the fixed effects estimator does.

To appreciate this point, let us first consider estimating Eq. (4.1) by the random effects model. Because the time-variant covariates $\mathbf{Z_{it}}$ and H_{it} (i.e. HIV

[9] One can, for example, decompose the time-invariant and time-variant variables in Eq. (4.1) as follows. Let $\mathbf{X}_{i1}$ and $\mathbf{X}_{i2}$ be respectively exogenous and endogenous time-invariant variables, and $\mathbf{Z}_{it1}$ and $\mathbf{Z}_{it2}$ be respectively exogenous time-variant and endogenous time-variant variables. The proposed instruments for endogenous variables at time t are $(\ddot{\mathbf{Z}}_{it1,}, \ddot{\mathbf{Z}}_{it2,}, \mathbf{X}_{i1}, \bar{\mathbf{Z}}_{i1})$, where $\ddot{\mathbf{Z}}_{it1,2} = \mathbf{Z}_{it1,2} - \bar{\mathbf{Z}}_{i1,2}$ and $\bar{\mathbf{Z}}_{i1,2} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{Z}_{it1,2}$. See, for example, Wooldridge (2002) for the detailed Hausman-Taylor procedure.

incidence) have both between-country and within-country variation, the random effects estimator can consistently estimate the effects of $\mathbf{Z_{it}}$ on HIV incidence if the within effects represented by $\boldsymbol{\theta}_{within}$ are equal to the between effects represented by $\boldsymbol{\theta}_{between}$. This condition, unfortunately, is easily violated (Zorn 2001). If, in fact, $\boldsymbol{\theta}_{within} \neq \boldsymbol{\theta}_{between}$ is true, the random effects approach would produce a vector of estimated coefficients $\hat{\boldsymbol{\theta}}$ that is equal to the weighted average of the within effects (i.e. $\hat{\boldsymbol{\theta}}_{within}$) and the between effects (i.e. $\hat{\boldsymbol{\theta}}_{between}$). However, this weighted average effect would not correctly reflect the effect of $\mathbf{Z_{it}}$ on HIV incidence, which is represented only by $\boldsymbol{\theta}_{within}$.

Unlike the random effects estimator, the within-between estimator avoids the restriction of $\boldsymbol{\theta}_{within} = \boldsymbol{\theta}_{between}$. This is because it decomposes $\boldsymbol{\theta}$ into within and between effects by re-expressing Eq. (4.1) as

$$H_{it} = \alpha + \beta R_i + \ddot{\mathbf{Z}}'_{it}\boldsymbol{\theta}_{within} + \bar{\mathbf{Z}}'_i\boldsymbol{\theta}_{between} + \mathbf{X}'_i\boldsymbol{\gamma} + u_{it}, \tag{4.2}$$

where $\ddot{\mathbf{Z}}_{it} = \mathbf{Z}_{it} - \bar{\mathbf{Z}}_i$, $\bar{\mathbf{Z}}_i = \frac{1}{T}\sum_{t=1}^{T}\mathbf{Z}_{it}$, and $u_{it} = a_i + b_t + \varepsilon_{it}$, so that $\boldsymbol{\theta}_{within}$ is not restricted to be the same as $\boldsymbol{\theta}_{between}$.

Eq. (4.2) can be estimated by generalized least squares regression if we are willing to assume that the unobserved components in Eq. (4.2) are uncorrelated with the regressors. If $\boldsymbol{\theta}_{within} = \boldsymbol{\theta}_{between}$, Eq. (4.2) will boil down to the usual random effects model. Otherwise, $\boldsymbol{\theta}_{within}$ will capture the within effect of $\mathbf{Z_{it}}$ in the same way as estimating Eq. (4.1) with country fixed effects.

The advantage of the within-between model, expressed in Eq. (4.2), is that it addresses the problem of correlated random effects concerning the time-variant factors, namely $Cov(\mathbf{Z}_{it}, a_i) \neq 0$. This is because by decomposing $\mathbf{Z}_{it}$ into its within (i.e. $\ddot{\mathbf{Z}}_{it}$) and between (i.e. $\bar{\mathbf{Z}}_i$) components, $\ddot{\mathbf{Z}}_{it}$ is by construction uncorrelated with a_i; and although $\bar{\mathbf{Z}}_i$ is correlated with a_i, $\bar{\mathbf{Z}}_i$ is controlled for in Eq. (4.2). Thus, Eq. (4.2) enables us to estimate the effects of the time-variant factors such as $\mathbf{Z}_{it}$ consistently.

What we then have to do is to focus on estimating the effect of terrain ruggedness, which is time-invariant. Now, this effect could be confounded by other geographical

factors contained in a_i. Given that a_i cannot be purged by fixed effects, as this would eliminate the effect of terrain ruggedness as well, we control for a large set of time-invariant geographical controls, denoted by $\mathbf{X}_i$. It is important to point out that in practice, we cannot guarantee that upon conditioning on $\mathbf{X}_i$, terrain ruggedness would be exogenous. However, it turns out that our conclusion does not depend on whether these geographical factors are controlled for. In other words, whether we control for $\mathbf{X}_i$ does not matter for whether we observe if terrain ruggedness has an effect on HIV incidence.

4.4 Results and Discussion

4.4.1 Preliminary Discussion

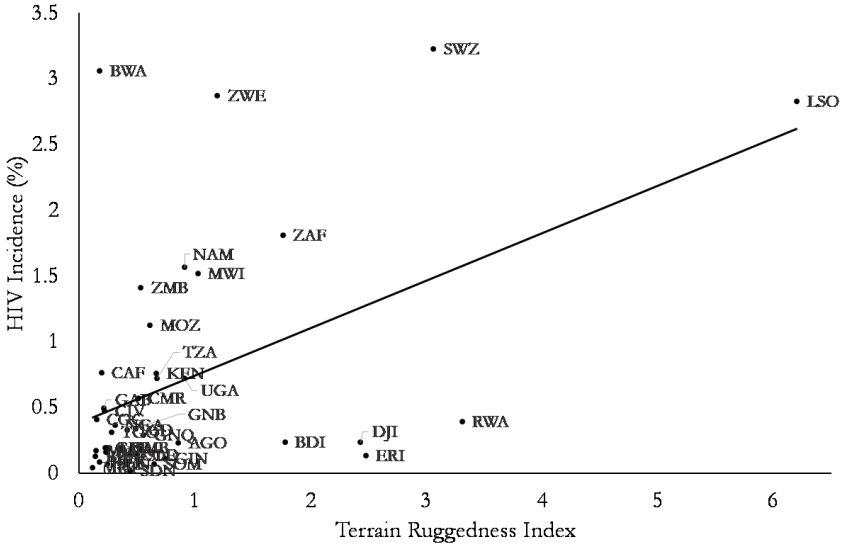
We first explore the correlation between the terrain ruggedness index and the average yearly HIV incidence rates between 1990 and 2014 in the SSAs. The scatter diagram in Figure 4.2 shows that there is a strong positive relationship between terrain ruggedness and HIV incidence. Spatially, there is a group of relatively highly rugged countries such as Swaziland and Zimbabwe that appear to have high HIV incidence rates, and most of these countries are located in the southern and eastern parts of Africa. Conversely, among countries in western and central Africa where their terrain is less rugged, the problem of HIV/AIDS appears to be less severe.

4.4.2 Baseline Estimates

We estimate the relationship between terrain ruggedness and HIV incidence using the random effects and within-between estimators, represented by Eqs. (4.1) and (4.2), respectively.

As a baseline, we include two time-variant controls, GDP per capita and population density (both in logs). It has been shown that in Africa, income and population density are correlated with ruggedness. For instance, countries with more rugged terrains tend to have higher incomes (Nunn and Puga 2012). Highly rugged places also tend to be less densely populated as rugged terrains are less suitable for

Figure 4.2: Terrain Ruggedness and HIV Incidence



farming and not easily accessible (Beckford and Barker 2007).

Income and population density may also be correlated with HIV incidence. Higher incomes, for instance, may encourage more commercial sexual activities as sex is a normal good (Ahlberg and Jensen 1998; Oster 2012).[10] Moreover, infectious diseases like HIV tend to spread faster in densely populated areas (Dyson 2003), especially in urban settings where interactions among individuals are more frequent (Kelly and van Donk 2009). This makes it important for us to control for both income and population density when estimating the effect of terrain ruggedness on HIV incidence.

The baseline results are reported in Table 4.2. The estimated coefficients on the terrain ruggedness index from the random effects (RE) and within-between (WB) estimators are both positive and statistically significant. This supports our hypothesis that higher levels of terrain ruggedness are associated with higher HIV incidence.

Other Remarks The within-between approach produces a larger coefficient on terrain ruggedness than what the random effects model estimates. For example, the random effects model shows that a one-standard deviation increase (which is about 120 meters) in terrain ruggedness is associated with a 0.29 percentage point increase in

[10]It is still empirically debatable whether income and risky sex are positively related. For example, Burke, Gong and Jones (2015) suggest otherwise that it is income decline that contributes to more HIV infections because in time of income crisis, young women are forced to engage in risky transactional sex usually with older men in exchange for economic stability.

Table 4.2: Effect of Ruggedness on HIV: Baseline Estimates

Dependent variable: HIV Incidence (1990-2014)	(1) RE	(2) WB
Ruggedness	0.244**	0.358***
	(0.112)	(0.082)
Log GDP per capita[a]	0.381***	0.220**
	(0.121)	(0.106)
Log population density[a]	0.246	2.612**
	(0.172)	(1.241)
Constant	−2.660	−0.213
	(1.119)	(1.322)
R-squared	0.276	0.397
Number of observations	796	796
Number of countries	33	33

Note: Robust standard errors are in parentheses. ***, **, * Significance level at 1%, 5%, and 10%, respectively. [a]: Within-effects estimates for within-between estimator.

HIV incidence on average (Column (1)). By contrast, the within-between approach shows that HIV incidence increases by 0.43 percentage point on average (Column (2)). This effect is rather large, considering that it is about 53% of the average HIV incidence rate in sub-Saharan Africa.[11]

Concerning the time-variant controls (i.e. income and population density), the results in Column (2) suggest that these are positively associated with HIV incidence rate, which is consistent with what the literature has reported.

4.4.3 With Time-Invariant Geographical Controls

In our baseline regressions, we find that terrain ruggedness is positively associated with HIV incidence. We now investigate if this observation is driven by permanent cross-country differences that are not controlled for. Because the within-between estimator does not eliminate the time-invariant unobservables as country fixed effects do (otherwise, the effect of terrain ruggedness would be eliminated as well), we now include several potentially important time-invariant correlates of terrain ruggedness as controls.

Soil Quality In Africa, there is evidence that areas with rugged terrains are endowed with better soil quality than areas with flat terrains (Nunn and Puga 2012).

[11]The average HIV incidence rate among the SSAs is 0.81%.

If fertile soil is linked to higher agricultural productivity, and thus income, this factor could be relevant for HIV incidence as income may influence HIV transmission (Oster 2012; Burke, Gong and Jones 2015).

In Column (1) of Table 4.3, we control for soil quality (i.e. proportion of land area with fertile soil). The estimates in Column (1) of Table 4.3 show that soil quality is positively and statistically significantly associated with HIV incidence, which suggests that HIV spreads more in countries with more fertile soil. However, despite including soil quality as a control, the effect of terrain ruggedness on HIV incidence remains positive and statistically significant at the 1% level. Interestingly, the coefficient on terrain ruggedness has the same magnitude as the coefficient reported in the baseline regression (Column (2) of Table 4.2), which suggests that the impact of ruggedness on HIV incidence is not confounded by whether soil quality is controlled for.

Diamond Production Mineral resources are potential correlates of terrain ruggedness or the rate at which HIV transmits. For instance, certain types of minerals, such as diamonds, are found mainly in southern Africa, which is highly rugged. Furthermore, natural resource abundance may affect income[12] and give rise to resource occupation and conflict (Silberfein 2004), which in turn may affect the spread of HIV/AIDS (UNAIDS 2003; Burke, Gong and Jones 2015). In addition, mining boom following the discovery of minerals such as diamonds and golds may attract migrants to the mining industry, and these migrants tend to be at higher risk of contracting HIV (Corno and de Walque 2012).

To control for the effects of mineral resources, we control for diamond production (i.e. the total carats of gem-quality diamond extracted from 1958 to 2000 normalized by country size). As Column (2) shows, diamond production has a positive impact on HIV incidence, as we would expect. However, it does not drive out the effect of terrain ruggedness on HIV, which remains positive and statistically significant at the 1% level.

[12]For example, there is evidence of resource-rich countries failing to achieve growth, experiencing a recession due to crowding out effects (i.e. natural resource dependence drives out other sectors like exports or services that contribute to growth) (see *inter alia*, Sachs and Warner 1999, 2001; Wick and Bulte 2006).

Table 4.3: Robustness Results Using Geographical Controls

	Dependent variable: HIV Incidence (1990-2014)				
Method: WB	(1)	(2)	(3)	(4)	(5)
Ruggedness	0.358***	0.372***	0.213**	0.363***	0.291***
	(0.079)	(0.080)	(0.101)	(0.080)	(0.098)
Log GDP per capita[a]	0.220**	0.221**	0.221**	0.220**	0.221**
	(0.106)	(0.106)	(0.106)	(0.106)	(0.106)
Log population density[a]	2.612**	2.612**	2.612**	2.612**	2.611**
	(1.242)	(1.242)	(1.242)	(1.242)	(1.245)
Soil quality	1.322**				1.438**
	(0.635)				(0.600)
Diamonds		0.281*			0.314***
		(0.163)			(0.075)
Tropical climate			−0.703*		−0.425
			(0.371)		(0.355)
Distance to coast				0.406	0.402
				(0.393)	(0.380)
Constant	−0.055	−0.054	−0.176	−0.966	−0.589
	(1.203)	(1.306)	(1.365)	(1.539)	(1.366)
R-squared	0.454	0.434	0.453	0.411	0.546
Number of observations	796	796	796	796	796
Number of countries	33	33	33	33	33

Note: Robust standard errors are in parentheses. ***, **, * Significance level at 1%, 5%, and 10%, respectively. [a]: Within-effects estimates.

Tropical Climate Rugged areas tend to have cooler climatic conditions compared with areas that are flat. Countries whose large proportion of land surface is covered by tropical climate tend to be more prone to diseases like malaria or parasitic infections (Mabey et al. 2004; Hay et al. 2005). This could lead to more HIV infections because firstly, disease burden in a household may trigger poverty, which may drive a female family member to engage in transactional sex for income (Burke, Gong and Jones 2015), and secondly, tropical diseases may increase the risk of HIV transmission, for instance, by causing malaria patients to contract HIV through blood transfusion (Morrow, Colebunders and Chin 1989) and by causing women with a parasitic disease to be exposed to HIV infections through genital tract infections.[13]

In Column (3), we control for tropical climate. We find that tropical climate is negatively associated with HIV incidence, although this association is rather weak (the coefficient on tropical climate is statistically significant at the 10% level). While controlling for tropical climate reduces the effect of terrain ruggedness on

[13]Parasitic disease could cause the genital tract of women to become infected, which consequently allows HIV to transmit more easily during intercourse (Mbabazi et al. 2011).

HIV incidence, this effect remains statistically significant. This suggests that terrain ruggedness has an impact on HIV incidence independently of the effect of tropical climate.

Distance to Coast Another possible confounding geographical factor is the average distance to the closest coast for each country. The relation between distance to coast and HIV may be explained by economic factors. In the literature, there is evidence that the coastal areas of countries tend to be economically more affluent than the interior areas are (Rappaport and Sachs 2003). In these areas, close proximity to sea and seaports may contribute to rising trade, productivity and therefore income. Trade and income in turn are associated with the spread of HIV (Oster 2012; Lin and Sim 2015). However, as Column (4) shows, the effect of distance to coast on HIV incidence is statistically insignificant. Not surprisingly, after controlling for it, the impact of terrain ruggedness on HIV is still positive and statistically significant with a magnitude that is close to the baseline estimate.

All Geographical Controls Finally, in Column (5), we control for all geographical factors at once and find that the effect of terrain ruggedness on HIV incidence is positive and statistically significant at the 1% level. Therefore, it is unlikely that our main result is confounded by omitted time-invariant geographical variables.

4.4.4 Robustness Checks

We perform three robustness checks here. The first robustness check looks at the sensitivity of our result to using alternative measures of the terrain ruggedness index and alternative data on HIV incidence. The second robustness check attempts to address the issue of model uncertainty by averaging across models using the Bayesian Model Averaging approach to estimate the effect of terrain ruggedness on HIV incidence. The final robustness check examines the sensitivity of our baseline result (i.e. terrain ruggedness has a positive effect on HIV incidence) to omitted time-variant factors, non-geographically related factors such as conflict, institution and infrastructure quality.

A. Alternative Data

We first examine how robust our main conclusion is to using alternative measures of the terrain ruggedness index and alternative data on HIV incidence. Firstly, to check if our main conclusion is robust to different measures of terrain ruggedness, we follow Nunn and Puga (2012) to consider four variables, namely, the population-weighted terrain ruggedness index, average absolute slope, standard deviation of elevation, and rate of highly rugged terrain (see Section 4.2 for their description). As Figure 4.3 (Panels A-D) shows, each of these alternative measures of the terrain ruggedness index is strongly and positively correlated with HIV incidence as well.

As such, in place of the terrain ruggedness index, we use each of alternative measures as an explanatory variable in the model while controlling for the geographical and time-variant covariates.[14] From Table 4.4 (Columns (1)-(4)), we find that each of these alternative measures has a positive and statistically significant effect on HIV (just as the terrain ruggedness index has). Thus, our main finding that terrain ruggedness is positively associated with HIV incidence in the SSAs is not sensitive to how ruggedness is measured.

Next, considering also the alternative measures of terrain ruggedness, we now examine how robust our results are when a different dataset on HIV incidence is used. Here, we reconstruct the HIV incidence data using a method suggested by Oster (2012). As discussed in Section 4.2, this method allows us to estimate the HIV incidence rates for the years before 1990 during which there is no data from UNAIDS. Thus, the alternative HIV dataset used here contains the HIV incidence rates from 1985 to 2014.

As Columns (1)-(5) of Table 4.5 show, terrain ruggedness and its alternative measures are all statistically significant at the 1% level for HIV incidence, when based on the alternative HIV incidence dataset. Therefore, our conclusion is not only robust to using the different measures of terrain ruggedness, it is also robust to using the alternative HIV incidence dataset that extends back to 1985.

[14] We exclude the variables which control for conflicts, quality of institution and infrastructure to avoid the loss of number of observations. However, including the full sets of control variables does not substantially alter our robustness results (see Appendix A).

Figure 4.3: Alternative Measures of Ruggedness and HIV Incidence

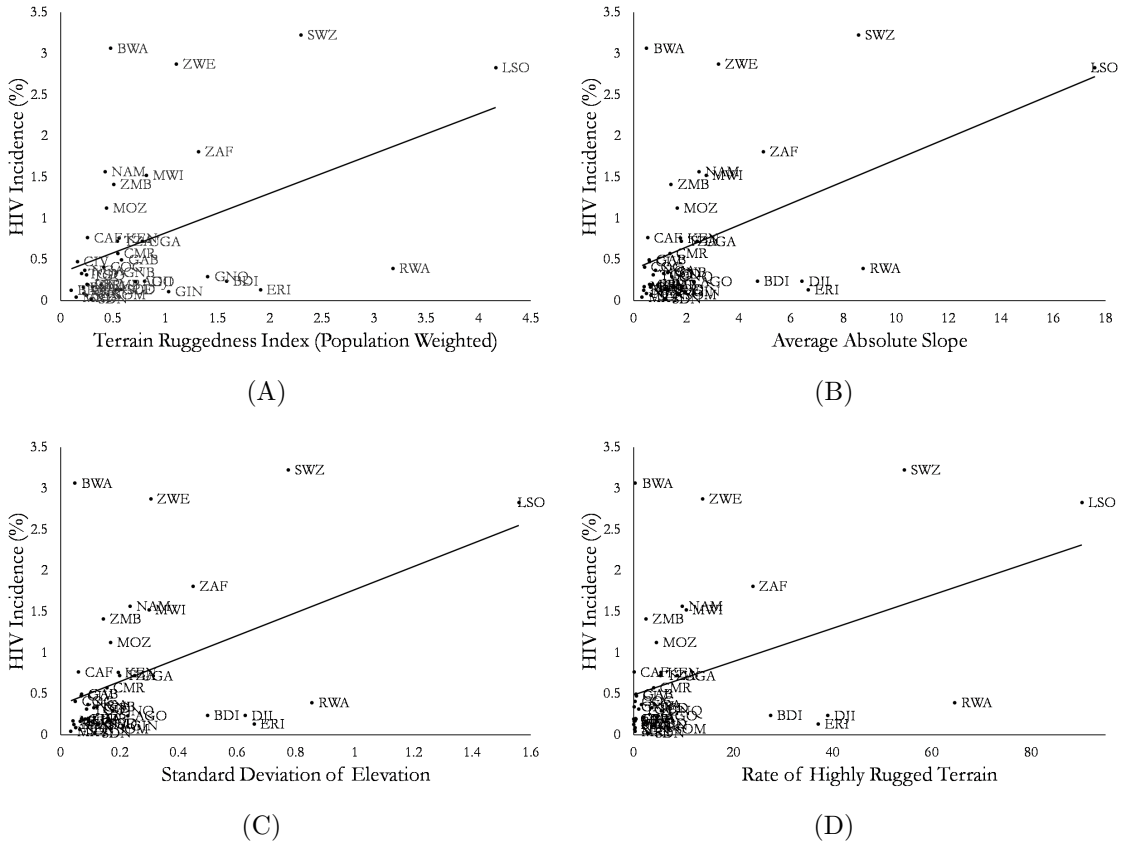


Table 4.4: Effect of Alternative Measures of Ruggedness on HIV

	Dependent variable: HIV Incidence (1990-2014)			
Method: WB	(1)	(2)	(3)	(4)
Ruggedness (population-weighted)	0.360**			
	(0.146)			
Average absolute slope		0.106***		
		(0.034)		
Standard deviation of elevation			1.171**	
			(0.388)	
Rate of highly rugged terrain				0.015*
				(0.008)
Constant	−0.217	−0.584	−0.604	−0.333
	(1.410)	(1.355)	(1.376)	(1.326)
Time-variant controls	Yes	Yes	Yes	Yes
Geographical controls	Yes	Yes	Yes	Yes
R-squared	0.541	0.549	0.547	0.526
Number of observations	796	796	796	796
Number of countries	33	33	33	33

Note: Robust standard errors are in parentheses. ***, **, * Significance level at 1%, 5%, and 10%, respectively. Time-variant controls include log of GDP per capita and log of population density. Geographical controls are soil quality, diamonds, tropical climate and distance to coast.

Table 4.5: Effect of Ruggedness on HIV Using Alternative Data on HIV

	Dependent variable: HIV Incidence (1985-2014)				
Method: WB	(1)	(2)	(3)	(4)	(5)
Ruggedness	0.311***				
	(0.091)				
Ruggedness (population-weighted)		0.369***			
		(0.131)			
Average absolute slope			0.113***		
			(0.031)		
Standard deviation of elevation				1.246***	
				(0.358)	
Rate of highly rugged terrain					0.016**
					(0.007)
Constant	−1.469	−1.227	−1.459	−1.500	−1.166
	(1.175)	(1.238)	(1.162)	(1.185)	(1.143)
Time-variant controls	Yes	Yes	Yes	Yes	Yes
Geographical controls	Yes	Yes	Yes	Yes	Yes
R-squared	0.545	0.536	0.548	0.546	0.524
Number of observations	896	896	896	896	896
Number of countries	37	37	37	37	37

Note: Robust standard errors are in parentheses. ***, **, * Significance level at 1%, 5%, and 10%, respectively. Time-variant controls include log of GDP per capita and log of population density. Geographical controls are soil quality, diamonds, tropical climate and distance to coast.

B. Bayesian Model Averaging

To investigate the robustness of our results, we have included or excluded covariates that are possibly relevant for the transmission of HIV/AIDS. Because there are different combinations in the choice of controls, this implicitly suggests that our model is itself uncertain.

To address the concern that many combinations of control variables can be used when estimating the impact of terrain ruggedness, which leads to model uncertainty as such, we estimate Eq. (4.2) through the Bayesian Model Averaging (BMA) approach. The BMA approach allows for a large set of models with different combinations of control variables to be estimated. It provides the weighted average of the estimates from those models based on the posterior inclusion probability (PIP) as weights (see *inter alia*, Hoeting et al. 1999; Fernández, Ley and Steel 2001; Moral-Benito 2012). The PIP is computed as the sum of all posterior probabilities related to each estimated model, and measures the likelihood that each variable belongs to the model. A variable with a PIP score that is close to 1 is a variable that almost certainly belongs to the model.

Table 4.6: Effect of Ruggedness on HIV: BMA Results

Dependent variable: HIV Incidence (1990-2014)	(1) WB	(2) PIP
Ruggedness	0.269**	1.000
	(0.115)	
Log GDP per capita[a]	0.202***	0.960
	(0.071)	
Log population density[a]	2.653***	1.000
	(0.330)	
Soil Quality	0.234	0.210
	(0.531)	
Diamond	0.140	0.400
	(0.196)	
Tropical climate	−0.426	0.550
	(0.445)	
Distance to coast	0.015	0.050
	(0.115)	
Constant	190.385	1.000
	(222.159)	
Number of observations	796	
Number of estimated models	256	

Note: Standard errors are in parentheses. ***, **, * Significance level at 1%, 5%, and 10%, respectively. [a]: Within-effects estimates for within-between estimator.

Table 4.6 presents the regression results using the BMA approach. Column (1) provides the weighted average of the estimates of terrain ruggedness and the time-variant and time-invariant (geographical) controls.[15] We find that terrain ruggedness still maintains its impact on HIV incidence. The weighted average of the estimated coefficient on terrain ruggedness is positive and statistically significant at the 1% level. This suggests that the effect of terrain ruggedness on HIV is robust to model selection. In Column (2), we provide the PIP score for each variable. We find that GDP per capita and population density have the highest PIP score of 0.96 and 1, respectively. This affirms our inclusion of GDP per capita and population density as baseline controls throughout our analysis.

C. Other Potential Confounders

We may be concerned that the effect of terrain ruggedness on HIV is confounded by other factors that have so far not been controlled for. To address this, we introduce control variables that are related to conflict, quality of institution and infrastructure,

[15]We avoid including time-variant control variables on conflicts, quality of institution and infrastructure as it would reduce the number of observations by more than 70%.

which in turn, could be correlated with both terrain ruggedness and HIV incidence.

Wars and Conflicts First, terrain ruggedness could be associated with wars and conflicts, although the direction of their relationship is ambiguous. For instance, terrain ruggedness may increase the likelihood of a conflict because a rough topography like mountains, cliffs or caves could be military strategic locations that provide protection against ambushes (Buhaug and Gates 2002; Fearon and Laitin 2003). However, rugged areas also make it difficult to mobilize armies and transport artilleries, and this may discourage insurgencies and conflicts. The relationship between conflicts and HIV is not entirely obvious as well. There is evidence of an increasing number of HIV infections in conflict-affected areas due to sexual violence and livelihood collapse which forces women to become sex workers (UNAIDS 2003). However, conflicts may in contrast decelerate HIV infections as a result of forced displacement of people and limited sexual interactions (Spiegel et al. 2007).

In any case, it is important to check if the effect of terrain ruggedness on HIV incidence observed in the baseline is driven by conflict. To do so, we introduce the four conflict variables – any war, intra-state conflict, inter-state conflict and conflict intensity – as controls. The results reported in Columns (1)-(4) of Table 4.7, where we include the four conflict variables into the baseline equations together with all geographical and time-variant controls, suggest that conflicts are associated with higher rates of HIV infections. Yet, despite including the conflict variables, the effect of terrain ruggedness on HIV remains positive and statistically significant. Therefore, terrain ruggedness has explanatory power on HIV incidence independently of the effect of conflict.

Institutional Quality Next, we check if the impact of terrain ruggedness on HIV is driven by the quality of institutions. Although it is unclear how terrain ruggedness and institution are related, existing literature reveals that geography affects institutions (Gallup, Sachs and Mellinger 1999; Acemoglu, Johnson and

Table 4.7: Robustness Results Using Time-Variant Controls

	Dependent variable: HIV Incidence (1990-2014)								
Method: WB	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Ruggedness	0.300***	0.274***	0.280**	0.311***	0.378***	0.379***	0.382***	0.380***	0.342***
	(0.087)	(0.088)	(0.112)	(0.088)	(0.092)	(0.086)	(0.094)	(0.093)	(0.094)
Any war[a]	0.140***								
	(0.052)								
Intra-state conflict[a]		0.162***							
		(0.059)							
Inter-state conflict[a]			0.057						
			(0.081)						
Conflict intensity[a]				0.093***					
				(0.034)					
Rule of law[a]					−0.013				
					(0.084)				
Voice and accountability[a]						−0.080			
						(0.083)			
Regulatory quality[a]							−0.063		
							(0.125)		
Government effectiveness[a]								0.009	
								(0.092)	
Paved roads									0.057
									(1.245)
Constant	−0.988	−1.142	−0.317	−0.890	−0.607	−0.394	−0.288	−0.577	−0.475
	(1.595)	(1.622)	(1.606)	(1.565)	(1.497)	(1.328)	(1.508)	(1.722)	(1.314)
Time-variant controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographical controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.580	0.579	0.554	0.580	0.620	0.614	0.611	0.616	0.612
Number of observations	542	542	542	542	510	510	510	510	696
Number of countries	33	33	33	33	33	33	33	33	29

Note: Robust standard errors are in parentheses. ***, **, * Significance level at 1%, 5%, and 10%, respectively. Time-variant controls include log of GDP per capita and log of population density. Geographical controls are soil quality, diamonds, tropical climate and distance to coast. [a]: Within-effects estimates.

Robinson 2001; Easterly and Levine 2003).[16] If rugged areas are endowed with rich natural resources, the local elites may develop institutions that favor the interests of their own as well as their associates. Such extractive institutions may, in turn, affect policies including those concerning HIV/AIDS. In Columns (5)-(8), we control for the quality of institutions using four variables: rule of law, voice and accountability, regulatory quality, and government effectiveness. We find no evidence that the quality of institutions has an impact on HIV incidence, and that adding these institutional quality controls into the baseline equation does not drive out the positive and statistically significant effect of terrain ruggedness on HIV incidence.

Infrastructure Quality Finally, ruggedness may influence the quality of physical infrastructure, which may affect the transmission of HIV/AIDS. One possibility is that the quality of infrastructure in highly rugged areas tends to be relatively poor, for the rocky or uneven terrain makes it hard or costly to build good roads or bridges. Lack of quality infrastructure may, in turn, hinder effort to mobilize resources to help combat HIV/AIDS. Moreover, poor accessibility may also restrict people's sexual interaction, and that could slow down the spread of the disease. To check if this is the case, we add a variable on paved roads, defined as the proportion of paved roads to total roads, to our baseline equation together with a full set of geographical and time-variant controls. As shown in Column 9 of Table 4.7, we find no evidence that road infrastructure affects HIV incidence nor it affects the sign and statistical significance of the impact of terrain ruggedness on HIV incidence.

4.4.5 Slave Trade as a Mechanism

We argue that historical slave trade is one mechanism that gives rise to the positive effect of terrain ruggedness on HIV incidence. Taking place between 1400 and 1900, slave trade is a unique historical event that has caused significant destruction to

[16]For example, Acemoglu, Johnson and Robinson (2001) find that during the period of colonization, Europeans adopted extractive institution in colonized areas where the rate of mortality was high.

Africa.[17] During this period, around 20 million slaves were captured, mostly by force, and shipped out of Africa to the Middle East, India, Europe, and America.

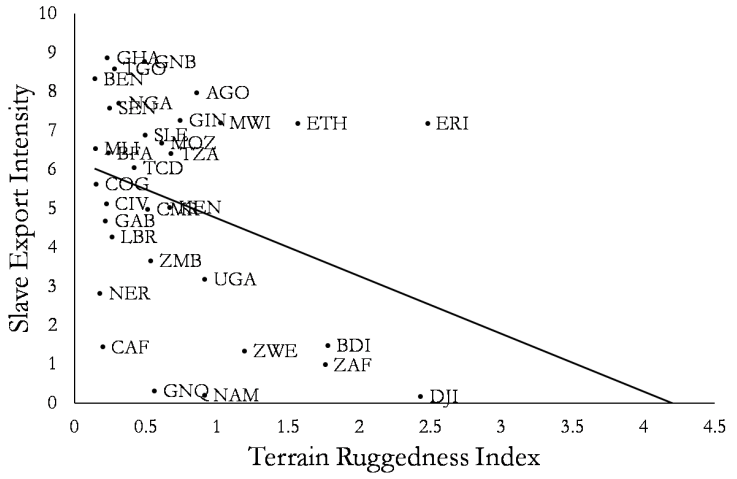
One reason for considering the slave trade as the mechanism for the effect of terrain ruggedness comes from the seminal work by Nunn and Puga (2012) who show that terrain ruggedness has a negative effect on slave export. The authors posit that rough topography like cliffs or caves provides protection from raids and subsequent enslavement. Rugged terrains also made it difficult for slaves to be transported to the nearest coast, which discourages slave capturing. Empirically, this is consistent with the fact that the large-scale slave export came mainly from western Africa, including countries such as Nigeria and Ghana that have relatively flat geography. By contrast, the number of slaves taken out of southern and eastern Africa, which are more rugged, was relatively small (see Figure 4.1, Panel B).

What is particularly interesting is that the spatial distribution of slave export appears to be in contrast with the spatial variation of HIV incidence in the SSAs. This suggests that slave export and HIV incidence could be correlated. For example, we present a scatter plot for the terrain ruggedness index and slave export intensity (see Section 4.2 for its description) and for slave export intensity and HIV incidence. The scatter plots in Figure 4.4 show that there is a strong negative association between terrain ruggedness and slave export intensity (Panel A), and as well a strong negative association between slave export intensity and HIV incidence (Panel B). This can also be shown more formally by regressing the terrain ruggedness index on slave export intensity using all geographical variables as controls, where Column (1) of Table 4.8 suggests that terrain ruggedness has a negative and statistically significant effect on slave export; and by regressing slave export intensity on HIV incidence, where Column (2) of Table 4.8 suggests that slave export has a negative association with HIV incidence.

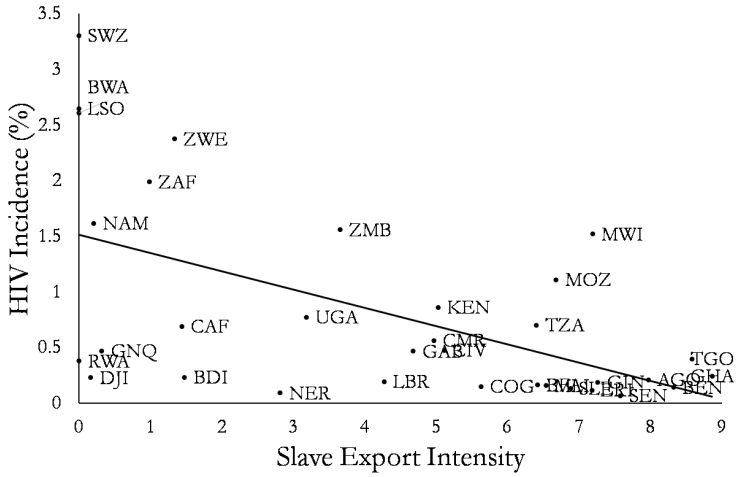
To check if the slave trade is the mechanism through which terrain ruggedness affects HIV incidence, we add slave export intensity to the baseline equation, which controls for the geographical and time-varying factors. As reported in Column (1)

[17]There is evidence of economic collapse as well as social and political divide as a result of the slave trade (M'baye 2006).

Figure 4.4: Terrain Ruggedness, Slave Trade and HIV



(A)



(B)

of Table 4.9, we find that slave export intensity is negatively associated with HIV incidence. However, it now *drives out* the once statistically significant effect of terrain ruggedness on HIV incidence. In other words, when slave export is controlled for, terrain ruggedness loses its explanatory power for HIV incidence. This suggests that slave export is the mechanism that underpins the effect of terrain ruggedness on HIV incidence.

To investigate how robust slave export is as the proposed mechanism, we consider using alternative measures of terrain ruggedness, just as in Section 4.4.4. Columns (2)-(5) of Table 4.9 show that slave export intensity, when it is added to the model

Table 4.8: Ruggedness, Slave Trade and HIV Incidence

Dependent variable:	Slave Export Intensity	HIV Incidence (1990-2014)
Method:	OLS	WB
	(1)	(2)
Ruggedness	−1.370***	
	(0.282)	
Slave export intensity		−0.146***
		(0.044)
Constant	4.899	1.428
	(1.768)	(1.114)
Time-variant controls	-	Yes
Geographical controls	Yes	Yes
R-squared	0.495	0.568
Number of observations	37	796
Number of countries	37	33

Note: Robust standard errors are in parentheses. ***, **, * Significance level at 1%, 5%, and 10%, respectively. Time-variant controls include log of GDP per capita and log of population density. Geographical controls are soil quality, diamonds, tropical climate and distance to coast.

(with various measures of terrain ruggedness), is negatively associated with HIV incidence. However, whereas the alternative measures of terrain ruggedness were previously statistically significant for HIV incidence (see Table 4.4), they all now become statistically insignificant once slave export is controlled for.

Similarly, when we repeat this exercise with the alternative data on HIV incidence (which extends the starting year of publicly available HIV incidence data from 1990 further back to 1985), we again find that terrain ruggedness, regardless of how it is measured, does not have any explanatory power for HIV incidence once slave export is controlled for (see Table A3 in Appendix A). This provides further evidence that slave export as the mechanism does not depend on how terrain ruggedness is measured and the HIV incidence dataset used.

Up to now, we have documented that countries with highly rugged terrain are likely to have exported fewer slaves in the past but have higher HIV incidence in the present. One explanation is that the association between the slave trade and HIV incidence is linked to trust. Firstly, Nunn and Wantchekon (2011) show that the slave trade has a persistent negative effect on trust. Slave trade in Africa was an unpleasant past, for slave capturing was not only done through violent attacks, but also by dirty tricks from close friends and relatives. To protect themselves, people in the community had to stay alert for possible raids or deceiving acts of others around

Table 4.9: Ruggedness and Alternative Measures, Slave Trade and HIV Incidence

Method: WB	Dependent variable: HIV Incidence (1990-2014)				
	(1)	(2)	(3)	(4)	(5)
Ruggedness	0.122				
	(0.099)				
Ruggedness (pop.-weighted)		0.149			
		(0.156)			
Average absolute slope			0.048		
			(0.033)		
Standard deviation of elevation				0.505	
				(0.397)	
Rate of highly rugged terrain					0.003
					(0.008)
Slave export intensity	−0.117***	−0.120***	−0.114**	−0.116**	−0.135***
	(0.045)	(0.045)	(0.046)	(0.046)	(0.041)
Constant	1.085	1.283	1.051	1.065	1.344
	(1.275)	(1.186)	(1.273)	(1.281)	(1.237)
Time-variant controls	Yes	Yes	Yes	Yes	Yes
Geographical controls	Yes	Yes	Yes	Yes	Yes
R-squared	0.577	0.577	0.579	0.577	0.569
Number of observations	796	796	796	796	796
Number of countries	33	33	33	33	33

Note: Robust standard errors are in parentheses. ***, **, * Significance level at 1%, 5%, and 10%, respectively. Time-variant controls include log of GDP per capita and log of population density. Geographical controls are soil quality, diamonds, tropical climate and distance to coast.

who might want to sell them into slavery. This kind of threatening neighborhood they lived in created a culture of mistrust which appear to persist till today (Nunn and Wantchekon 2011).

Secondly, trust is associated with risky sexual behavior that increases the spread of HIV. For example, in the epidemiology literature, there is evidence that people who display trust and faithfulness towards their sexual partners tend to forgo the use of protection during sex (Appleby, Miller and Rothspan 1999; Syvertsen et al. 2015). This appears to be relevant in the SSAs, where the use of sexual protection methods (i.e. condom use) in a trusting relationship is viewed as a sign of unfaithfulness and is therefore eschewed (Tavory and Swidler 2009; Coma 2014). Moreover, trust in sexual relationships induces individuals to ignore or underestimate the potential risks from having sex, including the risk of contracting sexually transmitted diseases like HIV. In addition, trust could also expand one's social network (**?**) including sexual network, and hence a pool of sexual partners, which increases the risk of infection (**?**; **?**). Therefore, in societies that are more trusting, HIV may spread more easily. This

may explain why terrain ruggedness is positively associated with HIV incidence in the SSAs, since the restriction of slave trade by terrain ruggedness historically may foster more trusting societies today.

4.5 Conclusion

This paper shows that the cross-sectional variation in HIV incidence is associated with a deep, fundamental factor such as geography. We exploit data on terrain ruggedness as a measure of geographical features of countries and show that rugged terrain a positive effect on HIV incidence. However, this positive association is driven away once we control for slave exports in the past. As such, the positive relationship between terrain ruggedness and HIV incidence is consistent with the following explanation. Firstly, terrain ruggedness has a negative association with the historical slave trade in Africa as rugged terrain restricts slave capturing. In turn, slave trade has a persistent negative effect on how people trust each other today. Given the empirical evidence that HIV/AIDS spreads more easily when there is trust, terrain ruggedness, which restricted slave trade in the past, may contribute to higher HIV incidence rates.

Our work highlights the complexity of HIV/AIDS epidemic in sub-Saharan Africa. We show that factors that influence risky sexual behavior of people leading to the spread of HIV/AIDS are not only confined to socio-economic characteristics, as the literature has documented. Rather, people's behavior including sexual attitudes could be intrinsic to deep-rooted factors pertaining to history and geography unique to certain countries or regions. This may explain to some extent why in some parts of Africa, HIV/AIDS has turned out to be a more serious issue than in others.

References

Acemoglu, Daron, Simon Johnson and James A. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91(5):1369–1401.

Ahlberg, Dennis and Eric Jensen. 1998. The Economics of the Commercial Sex Industry. In

Confronting AIDS: Evidence from the Developing World, ed. Martha Ainsworth, Lieve Fransen and Mead Over. Brussels: The European Commission pp. 147–174.

Alesina, Alberto, Paola Giuliano and Nathan Nunn. 2013. "On the Origins of Gender Roles: Women and the Plough." *Quarterly Journal of Economics* 128(2):469–530.

Allison, Paul D. 2009. *Fixed Effects Regression Models*. Thousand Oaks: Sage.

Appleby, Paul Robert, Lynn Carol Miller and Sadina Rothspan. 1999. "The Paradox of Trust for Male Couples: When Risking Is a Part of Loving." *Personal Relationships* 6(1):81–93.

Banerjee, Abhijit and Lakshmi Iyer. 2005. "History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India." *American Economic Review* 95(4):1190–1213.

Beckford, Clinton and David Barker. 2007. "The Role and Value of Local Knowledge in Jamaican Agriculture: Adaptation and Change in Small-Scale Farming." *Geographical Journal* 173(2):118–128.

Bell, Andrew and Kelvyn Jones. 2015. "Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data." *Political Science Research and Methods* 3(1):133–153.

Buhaug, Halvard and Scott Gates. 2002. "The Geography of Civil War." *Journal of Peace Research* 39(4):417–433.

Burke, Marshall, Erick Gong and Kelly Jones. 2015. "Income Shocks and HIV in Africa." *Economic Journal* 125(585):1157–1189.

Coma, Julia Cordero. 2014. "HIV Prevention and Marriage: Peer group Effects on Condom Use Acceptability in Rural Kenya." *Social Science and Medicine* 116(2014):169–177.

Corno, Lucia and Damien de Walque. 2012. "Mines, Migration and HIV/AIDS in Southern Africa." *Journal of African Economies* 21(3):465–498.

de Walque, Damien. 2007. "How Does the Impact of an HIV/AIDS Information Campaign Vary with Educational Attainment? Evidence from Rural Uganda." *Journal of Development Economics* 84(2):686–714.

de Walque, Damien. 2009. "Does Education Affect HIV Status? Evidence from Five African Countries." *World Bank Economic Review* 23(2):209–233.

Dieleman, Joseph L. and Tara Templin. 2014. "Random-Effects, Fixed-Effects and the Within-Between Specification for Clustered Data in Observational Health Studies: A Simulation Study." *PLoS ONE* 9(10):1–17.

Duflo, Esther, Pascaline Dupas and Michael Kremer. 2015. "Education, HIV, and Early Fertility: Experimental Evidence from Kenya." *American Economic Review* 105(9):2757–2597.

Dyson, Tim. 2003. "HIV/AIDS and Urbanization." *Population and Development Review* 29(3):427–442.

Easterly, William and Ross Levine. 2003. "Tropics, Germs, and Crops: How Endowments Influence Economic Development." *Journal of Monetary Economics* 50(1):3–39.

Fearon, James D. and David D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97(1):75–90.

Fernández, Carmen, Eduardo Ley and Mark F. J. Steel. 2001. "Model Uncertainty in Cross-Country Growth Regressions." *Journal of Applied Econometrics* 16(5):563–576.

Gallup, John Luke, Jeffrey D. Sachs and Andrew D. Mellinger. 1999. "Geography and Economic Development." *International Regional Science Review* 22(2):179–232.

Greene, William H. 2012. *Econometric Analysis*. Harlow: Pearson.

Hausman, Jerry A. and William E. Taylor. 1981. "Panel Data and Unobservable Individual Effects." *Econometrica* 49(6):1377–1398.

Hay, Simon I., Carlos A. Guerra, Andrew J. Tatem, Peter M. Atkinson and Robert W. Snow. 2005. "Urbanization, Malaria Transmission and Disease Burden in Africa." *Nature Reviews Microbiology* 3(1):81–90.

Hoeting, Jennifer A., David Madigan, Adrian E. Raftery and Chris T. Volinsky. 1999. "Bayesian Model Averaging: A Tutorial." *Statistical Science* 14(4):382–401.

Kelly, Kevin and Mirjam van Donk. 2009. Local-Level Responses to HIV/AIDS in South Africa. In *HIV/AIDS in South Africa 25 Years On: Psychosocial Perspectives*, ed. Poul Rohleder, Leslie Swartz, Seth C. Kalichman and Leickness Chisamu Simbayi. New York: Springer pp. 135–153.

Kremer, Michael. 1996. "Integrating Behavioral Choice into Epidemiological Models of AIDS." *Quarterly Journal of Economics* 111(2):549–573.

Lakdawalla, Darius, Neeraj Sood and Dana Goldman. 2006. "HIV Breakthroughs and Risky Sexual Behavior." *Quarterly Journal of Economics* 121(3):1063–1102.

Lin, Faqin and Nicholas C.S. Sim. 2015. "Exports, HIV Incidence and the Baltic Dry Index: Further Evidence from Sub-Saharan Africa." *Economics Letters* 126(Supplement C):35–39.

Mabey, David, Rosanna W. Peeling, Andrew Ustianowski and Mark D. Perkins. 2004. "Diagnostics for the Developing World." *Nature Reviews Microbiology* 2(2004):231–240.

Mbabazi, Pamela Sabina, Olivia Andan, Daniel W. Fitzgerald, Lester Chitsulo, Dirk Engels and Jennifer A. Downs. 2011. "Examining the Relationship Between Urogenital Schistosomiasis and HIV Infection." *PLOS Neglected Tropical Diseases* 5(12):1–8.

M'baye, Babacar. 2006. "The Economic, Political, and Social Impact of the Atlantic Slave Trade on Africa." *European Legacy* 11(6):607–622.

Moral-Benito, Enrique. 2012. "Determinants of Economic Growth: A Bayesian Panel Data Approach." *Review of Economics and Statistics* 94(2):566–579.

Morrow, Richard H., Robert L. Colebunders and James Chin. 1989. "Interactions of HIV Infection with Endemic Tropical Diseases." *AIDS* 3(1):S79–S88.

Nunn, Nathan. 2008. "The Long-Term Effects of Africa's Slave Trades." *Quarterly Journal of Economics* 123(1):139–176.

Nunn, Nathan and Diego Puga. 2012. "Ruggedness: The Blessing of Bad Geography in Africa." *Review of Economics and Statistics* 94(1):20–36.

Nunn, Nathan and Leonard Wantchekon. 2011. "The Slave Trade and the Origins of Mistrust in Africa." *American Economic Review* 101(7):3221–3252.

Oster, Emily. 2012. "Routes of Infection: Exports and HIV Incidence in Sub-Saharan Africa." *Journal of the European Economic Association* 10(5):1025–1058.

Plümper, Thomas and Vera E. Troeger. 2007. "Efficient Estimation of Time-Invariant and Rarely Changing Variables in Finite Sample Panel Analyses with Unit Fixed Effects." *Political Analysis* 15(2):124–139.

Rappaport, Jordan and Jeffrey D. Sachs. 2003. "The United States as a Coastal Nation." *Journal of Economic Growth* 8(1):5–46.

Riley, Shawn J., Stephen D. Degloria and Robert Elliot. 1999. "A Terrain Ruggedness Index that Quantifies Topographic Heterogeneity." *Intermountain Journal of Sciences* 5(1):23–27.

Sachs, Jeffrey D., Andrew D. Mellinger and John L. Gallup. 2001. "Geography of Poverty and Wealth." *Scientific American* 284(3):71–75.

Sachs, Jeffrey D and Andrew M. Warner. 1999. "The Big Push, Natural Resource Booms and Growth." *Journal of Development Economics* 59(1):43–76.

Sachs, Jeffrey D. and Andrew M. Warner. 2001. "The Curse of Natural Resources." *European Economic Review* 45(4):827–838.

Silberfein, Marilyn. 2004. "The Geopolitics of Conflict and Diamonds in Sierra Leone." *Geopolitics* 9(1):213–241.

Smith, Jordan D. 2007. "Modern Marriage, Men's Extramarital Sex, and HIV Risk in Southeastern Nigeria." *American Journal of Public Health* 97(6):997–1005.

Spiegel, Paul B., Anne Rygaard Bennedsen, Johanna Claass, Laurie Bruns, Njogu Patterson, Dieudonne Yiweza and Marian Schilperoord. 2007. "Prevalence of HIV Infection in Conflict-Affected and Displaced People in Seven Sub-Saharan African Countries: a Systematic Review." *Lancet* 369(9580):2187–2195.

Swidler, Ann and Susan Cotts Watkins. 2007. "Ties of Dependence: AIDS and Transactional Sex in Rural Malawi." *Studies in Family Planning* 38(3):147–162.

Syvertsen, Jennifer L., Angela Robertson Bazzi, Gustavo Martinez, Gudelia M. Rangel, Monica D. Ulibarri, Kirkpatrick B. Fergus, Hortensia Amaro and Steffanie A. Strathdee. 2015. "Love, Trust, and HIV Risk Among Female Sex Workers and Their Intimate Male Partners." *American Journal of Public Health* 105(8):1667–1674.

Tavory, Iddo and Ann Swidler. 2009. "Condom Semiotics: Meaning and Condom Use in Rural Malawi." *American Sociological Review* 74(2):171–189.

UNAIDS. 2003. HIV/AIDS and Conflict. UNAIDS Fact Sheet.

UNAIDS. 2010. *Report on the Global AIDS Epidemic 2010*. Geneva: Joint United Nations Programme on HIV/AIDS.

UNAIDS. 2016. "AIDSinfo." http://aidsinfo.unaids.org/. (accessed 10 January 2017).

Wick, Katharina and Erwin H. Bulte. 2006. "Contesting Resources: Rent Seeking, Conflict and the Natural Resource Curse." *Public Choice* 128(3):457–476.

Wooldridge, Jefrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Massachusetts: MIT Press.

Zorn, Christopher. 2001. "Estimating Between- and Within-Cluster Covariate Effects, with an Application to Models of International Disputes." *International Interactions* 27(4):433–445.

Appendix A: Supplementary Results

In this appendix, we further show that our main results are robust to controlling for all control variables and to the use of the alternative data on HIV incidence reconstructed using the method of Oster (2012). In the full set of control variables, in addition to log of GDP per capita, log of population density and geographical variables, we include variables on any war, rule of law and paved roads. In particular, Table A1 presents the robustness results using the full set of control variables; Table A2 also reports the robustness estimates using all controls, but with the alternative dataset on HIV incidence; and Table A3 presents the results on the slave trade as a mechanisms using the alternative dataset on HIV incidence.

Table A1: Robustness Results Using All Controls

	Dependent variable: HIV Incidence (1990-2014)			
Method: WB	(1)	(2)	(3)	(4)
Ruggedness (population-weighted)	0.535***			
	(0.107)			
Average absolute slope		0.149***		
		(0.023)		
Standard deviation of elevation			1.662***	
			(0.255)	
Rate of highly rugged terrain				0.024***
				(0.005)
Constant	−0.428	−0.964	−0.969	−0.417
	(1.555)	(1.421)	(1.427)	(1.412)
Time-variant controls	YES	YES	YES	YES
Geographical controls	YES	YES	YES	YES
Other controls	YES	YES	YES	YES
R-squared	0.789	0.804	0.802	0.795
Number of observations	225	225	225	225
Number of countries	29	29	29	29

Note: Robust standard errors are in parentheses. Significance level at 1%, 5%, and 10%, respectively. Time-variant controls include log of GDP per capita and log of population density. Geographical controls are soil quality, diamonds, tropical climate and distance to coast. Other controls include any war, rule of law and paved roads.

Table A2: Robustness Results Using All Controls and Alternative Data on HIV

	Dependent variable: Oster's HIV Incidence (1985-2014)			
Method: WB	(1)	(2)	(3)	(4)
Ruggedness (population-weighted)	0.514***			
	(0.086)			
Average absolute slope		0.136***		
		(0.023)		
Standard deviation of elevation			1.544***	
			(0.250)	
Rate of highly rugged terrain				0.022***
				(0.004)
Constant	−1.422	−1.729	−1.762	−1.212
	(1.315)	(1.217)	(1.215)	(1.214)
Time-variant controls	YES	YES	YES	YES
Geographical controls	YES	YES	YES	YES
Other controls	YES	YES	YES	YES
R-squared	0.820	0.820	0.822	0.814
Number of observations	249	249	249	249
Number of countries	32	32	32	32

Note: Robust standard errors are in parentheses. Significance level at 1%, 5%, and 10%, respectively. Time-variant controls include log of GDP per capita and log of population density. Geographical controls are soil quality, diamonds, tropical climate and distance to coast. Other controls include any war, rule of law and paved roads.

Table A3: Ruggedness, Slave Trade and HIV Using Alternative Data on HIV

	Dependent variable: Oster's HIV Incidence (1985-2014)				
Method: WB	(1)	(2)	(3)	(4)	(5)
Ruggedness	0.112				
	(0.111)				
Ruggedness (population-weighted)		0.127			
		(0.152)			
Average absolute slope			0.044		
			(0.038)		
Standard deviation of elevation				0.454	
				(0.444)	
Rate of highly rugged terrain					0.002
					(0.009)
Slave export intensity	−0.132***	−0.136***	−0.130**	−0.132**	−0.149***
	(0.047)	(0.043)	(0.047)	(0.048)	(0.047)
Constant	0.502	0.642	0.472	0.485	0.776
	(1.276)	(1.168)	(1.272)	(1.287)	(1.223)
Time-variant controls	YES	YES	YES	YES	YES
Geographical controls	YES	YES	YES	YES	YES
R-squared	0.595	0.594	0.597	0.595	0.588
Number of observations	896	896	896	896	896
Number of countries	37	37	37	37	37

Note: Robust standard errors are in parentheses. Significance level at 1%, 5%, and 10%, respectively. Time-variant controls include log of GDP per capita and log of population density. Geographical controls are soil quality, diamonds, tropical climate and distance to coast.

Chapter 5

Concluding Remarks

This dissertation examines the role of culture, history and geography in explaining the uneven spread of HIV/AIDS in sub-Saharan Africa. Our first empirical study in Chapter 3 investigates whether the transmission of HIV in sub-Saharan Africa is rooted in population genetic diversity. Using trade to provide the treatment in our difference-in-differences strategy, we show that trade causes HIV to propagate, yet the positive impact of trade on HIV is weaker in countries with high population genetic diversity, suggesting that population genetic diversity and HIV infections are negatively associated. Our empirical results further demonstrate that the channels through which population genetic diversity affects the transmission of HIV are genetic and cultural homophily, and trust. In societies whose populations are genetically similar, genetic and cultural homophily makes it easier for people to form sexual relationships and hence encourages them to have multiple partners. Genetic and cultural homophily also breeds high levels of trust which is found to have caused risky sexual behavior leading to the spread of HIV.

The second empirical study in Chapter 4 explores whether terrain ruggedness, a measure of the uneven geographical features of an area, has an impact on the transmission of HIV in sub-Saharan Africa. It has been shown that the problem of HIV/AIDS in sub-Saharan Africa is more severe in countries with relatively high terrain ruggedness. Our further investigation into the channel of impact reveals that terrain ruggedness has a negative effect on historic slave trade, which has consequences for how people trust each other today. Trust in turn is believed to

have propagated HIV infections. Rugged terrain helped protect those being raided during the slave trades as mountains and caves provided look outs and hiding places. Irregular or rough ground also made it hard to move slaves from one place to another. The slave trade affected the level of trust people had in one another, and that influence still resonates today. In short, historically, fewer people were captured and enslaved in countries with high terrain ruggedness; today, people in those countries tend to be more trusting of one another. High trust, as we have shown, is associated with high rates of HIV infections.

In conclusion, this dissertation shows that the roots of HIV infections in sub-Saharan Africa are beyond the socio-economic characteristics that previous literature has documented. We have demonstrated that HIV risk factors may also be related to culture, history and geography, the determinants that policy barely affects. Equally important, this thesis also sheds some light on a critical gap in research into the causes of HIV infections in sub-Saharan Africa, which is to address the question of why HIV/AIDS affects this sub-continent so disproportionally.