UNU-WIDER | 30 YEARS OF RESEARCH FOR DEVELOPMENT

WIDER Working Paper 2015/086

Does HIV/AIDS matter for economic growth in sub-Saharan Africa?

Anthony Mveyange, Christian Skovsgaard, and Tine Lesner*

September 2015

United Nations University World Institute for Development Economics Research

wider.unu.edu

Abstract: Estimating the impact of HIV/AIDS epidemic on economic growth is challenging because of endogeneity concerns. In this paper, we use novel data on male circumcision and distance from the first HIV outbreak as instrumental variables for the HIV/AIDS epidemic in 241 regions across 25 countries in sub-Saharan Africa during 2003–12. Our main finding shows that the impact of HIV/AIDS epidemic on economic growth is negative but statistically insignificant. Further investigation on the main channels through which HIV/AIDS may affect economic growth—namely human capital, population growth, and productivity—finds no impacts of the HIV/AIDS epidemic on these channels.

Keywords: HIV/AIDS, economic growth, endogeneity, sub-Saharan Africa
JEL classification: I130, O110, O55

Acknowledgements: I wish to thank UNU-WIDER for support during my PhD internship, where this study was developed further. We thank Peter Sandholt Jensen, Thomas Barnebeck Andersen, Mickael Bech, Paul Sharp, Mircea Trandafir, Philip Agger, Giovanni Mellace, and Morten Sodemann, Casper Worm Hansen, Battista Severgnini, Channing Arndt, Annika Lindskog, Arcangelo Dimico, and Jennifer Kwok. Thanks are also extended to the participants at the 2015 CSAE Conference; University California Santa Cruz Economics Seminar Series; 2012 and 2013 Danish Graduate Program in Economics (DGPE) workshop; and SDU-CBS 2013 and 2014 annual workshops. The usual caveats apply.

*All authors Department of Business and Economics, University of Southern Denmark; corresponding author: amveyange@gmail.com.

This study has been prepared within the UNU-WIDER PhD internship programme.

ISSN 1798-7237 ISBN 978-92-9230-975-6 https://doi.org/10.35188/UNU-WIDER/2015/975-6

Typescript prepared by authors.

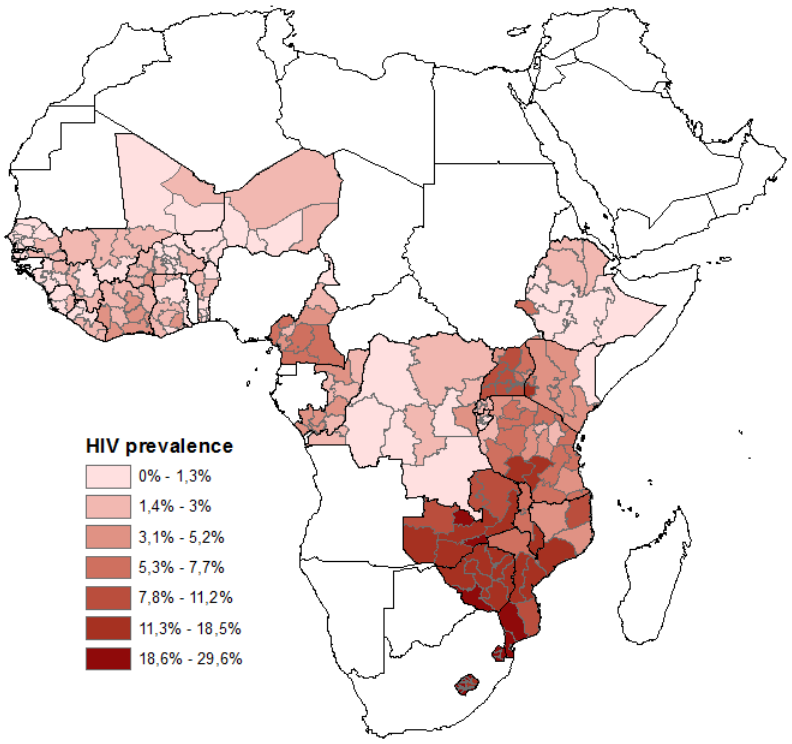
UNU-WIDER gratefully acknowledges the financial contributions to the research programme from the governments of Denmark, Finland, Sweden, and the United Kingdom.

UNU-WIDER, Katajanokanlaituri 6 B, 00160 Helsinki, Finland, wider.unu.edu

1 Introduction

Over the past three decades, the HIV/AIDS epidemic has been recognized as a major population health threat. Until 2012, 35 million people had died of the disease around the globe with sub-Saharan Africa recording a disproportionate share: 70 percent of all new global HIV infections take place in the region (UNAIDS, 2013). Figure 1 shows the visual spatial distribution of HIV prevalence rates by regional administrative units in 25 sub-Saharan African countries[1]. The HIV prevalence rates, as the figure shows, are largely concentrated in the so-called "AIDS belt" in the south-eastern part of Africa. The HIV epidemic has surely been a human disaster, but have the high prevalence rates undermined economic growth of the region? The goal of this paper is to provide fresh evidence on this widely debated question.

Figure 1: HIV Prevalence in Africa: 2003–2012



Source: Authors' construction based on DHS data

An abundant literature has tested the HIV – economic growth relationship. The overall agreement in the literature is that measures of HIV and economic development are correlated. However, the empirical evidence differs and provides mixed explanations on the causal association between HIV/AIDS and economic growth. As such, three broad strands of literature have emerged. The first strand asserts that the HIV/AIDS epidemic has negative effects on economic growth, mainly through its negative effects on health

[1] White shaded polygons are countries not covered in our sample.

budgets, labour supply, and human capital accumulation (see, for example, McDonald and Roberts, 2006; Alemu et al., 2005; Dixon et al., 2002; Bonnel, 2000; Cuddington, 1993a,b, for detailed discussions). The second strand points to the potential positive effects of the epidemic on economic growth indicators. This argument originates with Young (2005, p. 425) who argues that the "AIDS epidemic enhances the future per capita consumption possibilities" making it possible to achieve "higher standards of living for future generations" because the higher mortality and lower fertility increases the capital-labour ratio[2]. The third strand juxtaposes itself with the first two strands arguing that the epidemic has no significant effects on income per capita, (see, for instance, Werker et al., 2006; Bloom and Mahal, 1997). Several reasons have been attributed to this mixed evidence in the literature. For example, previous studies have mainly been plagued by the omitted variable bias that results from aggregating data at country level and endogeneity concerns between disease and development (see Bleakley, 2010; Ajakaiye and Audibert, 2010; Werker et al., 2006, for a detailed discussion).

This paper uses pre-epidemic male circumcision practices and the distance from the first epidemic outbreak as instruments for the HIV/AIDS epidemic to estimate its impacts on economic growth in sub-Saharan Africa during 2003–2012. The instrumental variables technique offers a credible way to establish the causal relationship between the HIV/AIDS epidemic and economic growth. The choice of male circumcision as an instrument[3] is plausible because male circumcision can reduce transmission risks: if not removed, the penis foreskin contains cells that makes it highly susceptible to HIV infection (see Donoval et al., 2006; Epstein, 2007; Timberg and Halperin, 2012, for details)[4]. Our circumcision instrument comes from Murdock et al. (1959) who provide data on ethnicities that practice male circumcision before the onset of the HIV epidemic in Africa.

The choice of distance from the reported first HIV outbreak (i.e. Kinshasa in the Democratic Republic of the Congo) as an instrument is based on epidemiological tenets on infectious diseases: a contagious disease outbreak is highly virulent to surrounding areas (Worobey et al., 2008). Moreover, we build on Oster (2012, p.36), who asserts that "in principle, if the virus takes time to travel, moving from person to person, areas further from its origin should have lower prevalence on average." The use of distance is naturally related to the potential pathways through which the spread of the disease

[2]However, more recent research has questioned this conclusion. Especially, Kalemli-Ozcan (2012) who finds that HIV does not lead to a fall in fertility, but instead an increase due to a precautionary demand for children as a response to the higher mortality risks.

[3]We are not the first to use male circumcision as an instrument. Werker et al. (2006) exploit the same idea using aggregated circumcision estimates from Halperin and Bailey (1999) but in a cross-country setting which renders their sample relatively small compared to ours.

[4]The evidence also shows that "HIV is less common among populations that traditionally practised male circumcision than in communities where the procedure is rare." See, `http://www.avert.org/male-circumcision.htm`.

ensued[5].

The main strength of our study, relative to previous literatures on the HIV/AIDS – economic growth debate, is that it focuses the analysis at sub-national level. This offers several advantages. First, it allows capturing and exploiting within and across countries the local scales variation of HIV/AIDS and its effects on economic growth. We argue that our approach offers more meaningful estimates than those normally offered by studies that use countries as their unit of analysis. Second, regional level data allow us to control for country-specific effects which naturally cannot be handled in a cross-country setting. As country fixed effects may be highly correlated both with economic growth and HIV/AIDS, the possibility to control for these in a regional set-up strengthen the analysis compared to previous cross-country studies. Third, the recent availability of spatial light data as a proxy for economic activity remedies the lack of consistent and reliable sub-national income data in Africa that has previously limited analysis at sub-national level.

We find that the impact of the HIV/AIDS epidemic on economic growth in sub-Saharan Africa is negative but statistically insignificant. Our results are robust across a range of specification tests. We also rule out other channels through which the epidemic can affect economic growth. We find insignificant effects of the epidemic on the three potential channels, namely, population growth, human capital and productivity. Neither do we find any significant effect of the epidemic on spatial inequality. However, we find our two instruments – male circumcision and distance to the first outbreak – as significant correlates of HIV-prevalence.

Our results relate to a recent body of literature investigating the effects of the HIV/AIDS epidemic and diseases in general on economic growth in Africa. For example, it resonates the recent evidence by Werker et al. (2006), who show that HIV/AIDS epidemic does not appear to affect economic growth in sub-Saharan Africa in recent years. Our findings are also in line with Acemoglu and Johnson (2007) and Soares (2007), who find little or no effect of health outcomes on economic indicators in the developing world.

The rest of the paper proceeds as follows. The next section provides a brief background on the HIV epidemic. Section 3 introduces the main data and their sources to familiarize the reader with our sub-national dataset. Section 4 presents the empirical strategy employed to establish the causal links. Section 5 presents and discusses our instrumental variables strategy that adresses inherent endoegeneity problems. Sections 6 and 7, respectively, present the main results and robustness checks. Section 8 discusses the results and concludes.

[5]In fact, it echoes Timberg and Halperin (2012)'s argument that trade and migration have been the main catalysts for the spread of the disease in Africa.

2 The HIV/AIDS Epidemic in Africa

In 1983, medical scientists in the United States discovered the HIV virus that causes AIDS. However, the virus existed long before that. Medical research based on blood plasma samples from Africans in 1959 shows that HIV was introduced into humans as early as around 1900 (Korber, 2000; Worobey et al., 2008). The general consensus in the medical literature is that HIV was transmitted to humans from chimpanzees. Cross-species transfer of the HIV virus to humans likely occurred because chimpanzees traditionally have been used as a source of livelihood in central Africa[6] (Timberg and Halperin, 2012).

How the virus spread is likely connected to the historical development[7] of Africa. New colonial cities with increasing population size and density, new trade routes and better infrastructure leading to an increase in travel activities paved the way for a global epidemic that since the 1990s has manifested itself as a serious threat to population health.

In sub-Saharan Africa the first HIV tests were performed and confirmed positive in the first half of the 1980s in the Democratic Republic of the Congo (DRC), Rwanda, and Burundi. Specifically, Kinshasa in the DRC has been reported as the epicentre of the HIV/AIDS epidemic and is documented to be the origin of the so-called Group M virus type that started the global epidemic (Worobey et al., 2008). Others, for example Oster (2012), argue that the precise location of the first outbreak within DRC is unknown and thus associate the origin as somewhat close to the centre of DRC.

The epidemic has affected sub-Saharan Africa disproportionately relative to other regions of the world: nearly 1 in 20 adults is infected with HIV (Haacker, 2004; UNAIDS, 2013). Over the last decade, however, there has been a decline of 33 percent in HIV incidence[8] which has resulted in a surge in HIV prevalence. We speculate that enhanced access[9] to antiretroviral treatment (ART) which prolongs the lives of the victims is the most likely explanation for these patterns.

[6]Two broad categories of HIV exist: HIV-1 and HIV-2. Because of its higher transmission risks and faster progression into AIDS relative to HIV-2, HIV-1 is the main cause of global infections and deaths (Worobey et al., 2008). HIV-2 accounts for approximately 5 percent of all HIV infections. The analysis of this paper does not specifically distinguish between HIV-1 and HIV-2.

[7]To quote Timberg and Halperin (2012), the "HIV/AIDS epidemic's birth and crucial early growth happened amid massive intrusion of new people and technology in Africa."

[8]The recent data on HIV prevalence rates show alarming increasing rates among adolescent and youth populations, especially females in sub-Saharan Africa UNAIDS (2013).

[9]Increased access to ARTs has also been beneficial in preventative terms since treatment reduces transmission risks of up to 96 percent (UNAIDS, 2013).

3 Data

In this section we describe the data used for the analysis. First, we present our data on night lights and explain how we use them as proxies for economic growth. Second, we describe our data on HIV prevalence whose main source is the geo-coded Demographic and Health Surveys (DHS) biomarker data. These data are nationally representative with $5,000$-$30,000$ households generally sampled in each country per survey round. Third, we describe the data that we use to construct and estimate our instruments. Fourth, we briefly describe other covariates that we extract from the DHS data.

3.1 Night Lights Data

Night lights (henceforth, lights) data have, in recent years gained traction in their use in empirical economics (see for example, Michalopoulos and Papaioannou, 2013; Henderson et al., 2012; Chen and Nordhaus, 2011)[10]. We use lights data to proxy for economic growth because of their proven high correlations with traditional output indicators and population settlement data at political jurisdiction level. Papageorgiou and Stoytcheva (2008) also show high correlations between light data and wealth indexes based on DHS data. Further, supported by Chen and Nordhaus (2011), Henderson et al. (2012) argue that these data are particularly valuable in a situation where output data are non-existent, a typical case for sub-national data for Africa[11].

The theoretical underpinning, when abstracting from public and private goods discussions alluded to by Henderson et al. (2012), of using light data draws on the tenet that consumption of lights at individual and household level is a function of income. Put differently, as people get richer the higher their consumption of lights becomes. Notwithstanding, the lights measured from outer space is not a perfect measure of economic activity. Michalopoulos and Papaioannou (2013) and Chen and Nordhaus (2011) describe saturations–when the high intensity of light is top-coded – particularly in rich countries, blooming – the high likelihood of observing dark place especially in poor countries – , and over-glow – lights being wrongly recorded as a reflection of another area – as key issues of concern. However, Michalopoulos and Papaioannou (2013) and Henderson et al. (2012) show that these are not pressing concerns for Africa. This strengthens the case for using lights as proxies for estimating economic growth in the present context[12]. Night lights have proven to be a good proxy especially for estimating

[10]The main source of these data is the US Defence Meteorological Satellite Program Operational Linescan System (DMSP-OLS) archived by the National Oceanic and Atmospheric Administration (NOAA). The data can be accessed and downloaded at `http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html`.

[11]Also problems like measurement errors, lack of consistency, missing data, manipulation – among others – that have tainted output data based on political jurisdiction can be overcome, to a greater extent, by using light data.

[12]In our sample the fraction of top-coding is less than 0.1 percent, with extremely trivial effect on our

economic *growth*[13].

We extract light data to cover the period 2003–2012. These data are normally measured daily from 8.30 p.m. to 10.00 p.m. local time. Annual averages are then estimated based on daily observations. The data are available on a 30 arc second resolution (corresponding to 0.86 square kilometre measured at the equator). The data values are coded on a scale of 0–63 with higher values reflecting more light density, i.e., more economic activity and vice versa for lower values. The data can either be obtained in a raw version or in a stable lights version adjusted for ephemeral activities such as forest fires, auroral activities, moonlight, and summer sun light. We use the stable lights data as our main source to estimate proxy for economic growth. Since these data include gas flares in some countries, we remove them before proceeding with the growth estimations.

The average growth rate of lights per capita is estimated between the year of the latest available DHS survey in each region[14] and country to the most recent year of light data availability, that is 2012. Spatial population data are extracted from the LandScan database[15].

The average light growth estimations are based on the following formula:[16]

$$g^{pc}_{i,j,sy,2012} = \frac{1}{N_i} \ln \left(\frac{Light_{i,j,2012}}{Population_{i,j,2012}} / \frac{Light_{i,j,sy}}{Population_{i,j,sy}} \right), \tag{1}$$

where, N is the number of years between the survey year and 2012, i is a country, j is a region, and sy stands for DHS survey year. $g^{pc}_{i,j,sy,2012}$ is annual average growth of per capita output for country i, region j from the survey year sy until 2012.

The use of light data offers yet another advantage: it allows us to estimate spatial inequality across regions in sub-Saharan Africa making it possible to test the effects of the epidemic on spatial inequality. Thus, we follow Alesina et al. (2015) to estimate spatial inequality across countries in our sample. Our measure of spatial inequality exploits grid level variations in light intensity to estimate spatial gini[17] for years between

results.

[13] In fact, as suggested by Pinkovskiy and Sala-i Martin (2014), the inherent and consistent processes that underlie the construction of light data have rendered them good proxies for estimating economic growth.

[14] We follow DHS definitions – which are based on political boundaries in a country – of regions for consistency. Note that these regions are referred to differently across countries – some are referred to as departments, districts, counties or provinces. However, in the DHS all these differences are harmonized to refer to regions.

[15] `http://web.ornl.gov/sci/landscan/landscan_data_avail.shtml`

[16] The formula for average growth rate in sum of light (a proxy for total growth) is given as: $g^{y}_{i,j,2012,sy} = \frac{1}{N_i} \ln \left(\frac{Light_{i,j,2012}}{Light_{i,j,sy}} \right)$.

[17] We use this formula:

$$Gini = \frac{\sum_{i=1}^{n} (2i - n - 1) * y_i}{n^2 P} \tag{2}$$

where i is grid cell rank order, n total number of grid cells, y_i is grid cell value lights per capita, and P is grid cell population average.

2003 and 2012.

3.2 HIV Prevalence Rates

Data on HIV prevalence rates are based on blood test results from DHS-surveyed individuals in sub-Saharan Africa. We use the blood test data to calculate regional-level HIV prevalence rates taking into account individual weights w_k given in the survey using the following formula[18]

$$HIV\,prevalence_j = \frac{\sum_{k \in j} HIV\,status_k \times w_k}{\sum_{k \in j} w_k}. \tag{3}$$

Equation 3 calculates the weighted HIV prevalence rate in region j by summing over the weighted outcomes of individuals k in region j and dividing this sum by the sum of weights within region j. Thus, we obtain representative regional prevalence rates for each country in our sample. We consider only surveyed individuals in the range of 15-49 years, which allows consistency across regions and gender.

Table 1 lists the DHS survey countries that we are able to include in our analysis. The table also shows when surveys were conducted in the countries and the number of regions in each country. Although the data consists of household surveys that are run across a large number of countries and over years, the data have a unbalanced panel structure. Thus, our primary analysis disregards the panel structure and instead uses a pooled cross-section structure of the data. We use countries with one survey round and for countries with multiple survey rounds we use the latest survey round[19]. Tables 17 and 18 in the appendices present additional empirical results based on pooling all the surveys from all countries.[20]

In Table 1, countries are sorted in descending order of their mean prevalence rates. The "AIDS belt" is represented by the top nine rows with Swaziland being the most and Tanzania being the least affected country in the "AIDS belt." The table shows considerable within-country variation, both for the more and the less affected countries. In Tanzania, the lowest regional prevalence rate is less than1 percent while the highest is almost 16 percent, whereas in Swaziland regional prevalence ranges between 23 and 29 percent. In Cameroon and Ethiopia the rates vary, respectively, from 1.2 to 7.2 percent and 0.9 to 6.5 percent. Certainly, these substantial within-country variations also justify our choice of sub-national level as a unit of analysis in estimate the economic effects of HIV/AIDS.

[18] Similar to Oster (2012) we also prefer to use HIV prevalence (*i.e. stock of infections*) rather than incidence rates (*i.e. new infections*). To get correct sub-national estimates and inferences, we follow the DHS manual and de-normalize weights when calculating totals c.f. equation 3.

[19] For Cote d'Ivoire we use the first survey since the latest survey is from 2012 – the same year as the latest light data. Hence, we would not be able to calculate any growth rates.

[20] Even after pooling the samples, our main results are unchanged.

Table 1: DHS Surveys and HIV Prevalence

Country	Survey Year	Number of Regions	HIV Prevalence (%)		
			Mean	Lowest	Highest
Swaziland	2006-07	4	25.88	23.18	28.87
Lesotho	2004, 2009	10	22.97	15.87	26.46
Zimbabwe	2005-06, 2010-11	10	15.22	13.36	21.19
Zambia	2001-02, 2007	9	14.28	6.79	20.77
Mozambique	2009	11	11.10	4.60	25.10
Malawi	2004, 2010	3	10.61	6.58	14.51
Uganda	2011	10	7.34	4.12	10.60
Kenya	2003, 2008-09	8	6.28	0.92	13.89
Tanzania	2003-04, 2007-08	21 (26)	5.68	0.08	15.72
Cameroon	2004, 2011	12	4.32	1.16	7.24
Cote d'Ivoire	2005, 2012	11	4.16	1.70	6.10
Congo	2009	12	3.16	1.46	4.84
Rwanda	2005, 2010	12 (5)	3.04	2.11	7.34
Ghana	2003	10	2.15	0.95	3.73
Liberia	2007	6	1.60	0.63	2.69
Sierra Leone	2008	4	1.51	0.83	2.89
Guinea	2005	8	1.50	0.73	2.13
Ethiopia	2005, 2011	11	1.47	0.87	6.50
Burundi	2010	5	1.41	0.86	3.71
Mali	2001, 2006	9	1.31	0.64	2.04
Dem. Rep. of Congo	2007	11	1.30	0.20	3.72
Benin	2006	12	1.15	0.30	2.70
Burkina Faso	2003, 2010	14 (13)	1.02	0.23	2.01
Niger	2006	8	0.74	0.35	1.72
Senegal	2005, 2010-11	11 (14)	0.67	0.15	2.40

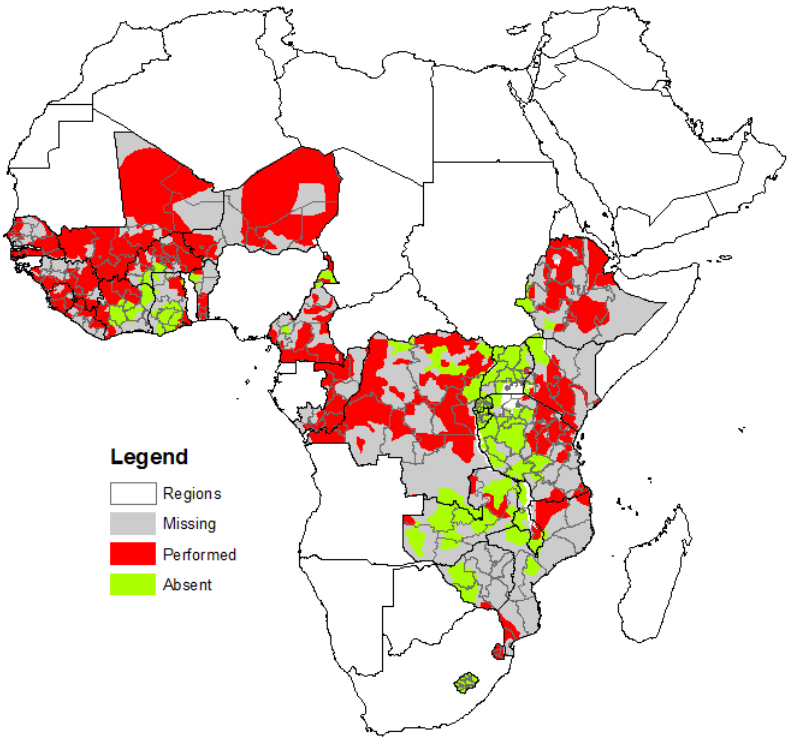
Notes: For Benin 2006 the country level prevalence rate is calculated as the simple average of the male and female prevalence rates due to lack of data. All prevalence rates are for men and women ages 15 – 49 years reported in percentages. The numbers in the parentheses refer to the number of surveyed regions in respective DHS years.

3.3 Male Circumcision

As previously mentioned, to construct our pre-determined male circumcision instrument we use Murdock et al. (1959)'s historical ethnographic data for more than 400 tribes in sub-Saharan Africa. Murdock et al. (1959) describe each tribe on a wide range of cultural parameters and in particular identify whether the tribes practice male circumcision as a tradition.

To estimate sub-national male circumcision shares on present day regions we cut the historical data by the present day regional borders and calculate the weighted share of the region in which male circumcision was traditionally performed. The construction of the male circumcision instrument is, however, limited by missing observations because of imperfect geographic coverage of the Murdock et al. (1959) data. To mitigate this limitation, similar in spirit to Werker et al. (2006), we assume 50 percent of males were traditionally circumcised in regions with missing data – an estimate close to the mean circumcision rate in sub-Saharan Africa. Figure 2 shows the ethnicity map of Murdock et al. (1959) within the boundaries of DHS's geographical coverage.

Figure 2: Male Circumcision Practices



Source: Authors' construction based on data from Murdock et al. (1959)

As a cross-check to our historical data we also compare the calculated circumcision rates with present day estimates from DHS[21]. The correlation is about 0.60, strong and significant.[22]

3.4 Distance

Our main distance measure is the population weighted estimate of average travel time in days from each country's region to Kinshasa in DRC. As partially alluded to before, the underlying conjecture of using this distance measure is that the spread of HIV virus depends on human interactions. Thus, the time (in days) it takes to travel from one place to another is a fairly good proxy for the rate of the epidemic spread. We estimate the number of days it takes to travel between Kinshasa and other regions based on data on land cover, roads, rails, trails, rivers and border frictions similar in spirit to

[21] We construct and calculate regional average male circumcision based on the DHS surveys. Filtering the data for the 15-49 years age range and non-respondents, we employ equation 3 to estimate regional shares of male circumcision. Response bias may likely affect our data, but we have no way of controlling for it. We refrain from using the present day male circumcision rates as instrument since they might be endogenous to growth especially as male circumcision campaigns have been performed post 2008. If campaigns for instance were systematically performed in poorer regions the instrument would be non-orthogonal to the unexplained part of economic growth.

[22] If we ignore areas with missing data the correlation becomes about 0.70, still strong and significant. These results are unreported but can be obtained from the authors upon request.

(Nelson, 2008). Since HIV is less likely to spread in an unpopulated areas we account for population densities when estimating the distance instrument[23].

To test the reliability[24] of our measure we also estimate an alternative measure of distance: the estimated travel time in days from each region to the centroid of DRC (coordinates: 6.31, 23.59) consistent with Oster (2012).[25]

3.5 Other Data

The DHS data further provide us with information on demographic characteristics. In our analysis we also extract two measures that we use to investigate potential channels through which the epidemic can affect economic growth: the average years of education (a proxy for human capital accumulation) and the share of people who were working (used to estimate the proxy for productivity – output per worker). The estimates of the shares of people working is based on a DHS question about whether they had been working for the previous year[26]. We then use the share of people working to estimate the proxy for the total output per worker[27]. Finally, the estimation of population growth is based on the LandScan dataset.

The Murdock et al. (1959) dataset also allows us to construct a measure of ethnicity. As we explain more in Section 5, controlling for ethnicity helps purge the threats to our identification strategy. However, constructing a measure of ethnicity in the presence of diverse tribal groups becomes a challenging task – it is difficult to code and properly account for all tribes. Thus, to get a sense of the ethnic composition we follow Michalopoulos and Papaioannou (2013)'s procedure to estimate an index of political centralization of ethnic groups by their geographical locations to estimate ethnic compositions within regions across countries.

[23] See the Appendix for more information on these estimations.

[24] We also calculate great circle distances in kilometres as an instruments. Regardless of the measure we use, the results remain qualitatively the same. Results can be obtained from the authors upon request.

[25] Oster (2012) proposes the middle of DRC in view of the argument by Sonnet et al. (1987) that early infections happened throughout the DRC.

[26] We also adjust for respondent answers that had been on leave for the last 7 days.

[27] Using the following formula: $\frac{Lights}{Population \times Share of working people}$

4 Empirical Strategy

To estimate the causal relationship between HIV/AIDS and economic growth we apply the 2SLS method with the following second-stage equation:[28]

$$\begin{aligned} g^{pc}_{i,j,sy,2012} = \gamma_0 + \gamma_1 Light pc_{i,j,sy} + \gamma_2 \hat{HIV}_{i,j,sy} \\ + X'_{i,j,sy}\gamma_3 + TE_{sy} + FE_i + \nu_{i,j,sy}, \end{aligned} \tag{4}$$

where $g^{pc}_{i,j,sy,2012}$ is regional light per capita growth (a proxy for economic growth). $\hat{HIV}_{i,j,sy}$ indexes the predicted average regional shares of HIV prevalence rates during the survey year in a country. $Light pc_{i,j,sy}$ denotes light per capita across regions and countries during the survey years. We use it as a proxy for initial income per capita which controls for any epidemic-economic growth confounding effects prior to countries' specific survey years. We also include country fixed effects, FE_i, to control for country-level unobserved characteristics.[29] Country fixed effects allow us to exploit within country variation – as noted before, previous studies relying on country level data failed to account for this. Further, we include time (survey year) fixed effects TE_{sy} to control for any time specific effects.

The vector $\mathbf{X_{i,j,sy}}$ includes mainly controls to the threat of our identification strategy. Section 5 elaborates on these controls in detail. The controls include regional shares of Muslim populations, and measures of ethnicity as in Michalopoulos and Papaioannou (2013). As Timberg and Halperin (2012) and Oster (2012) point out, the spread of HIV has been biased towards capital cities. That is, the HIV/AIDS epidemic arrives first and becomes most rampant in capital cities and then spreads to other regions. Therefore, we account for whether a region contains a capital city using a binary indicator variable. ν is an error term.

The main explanatory variable in our estimating equation is HIV prevalence rate. As such, our parameter of interest is coefficient γ_2. This coefficient captures the local average effects (LATE) of the epidemic on economic growth. The coefficient, while important in telling us the size of the effect, offers little insights for policy: if indeed causal effects exist, it fails to explain the mechanisms behind such an effect. Therefore, we subject our baseline specification to an implicit test: we estimate the causal effects of the epidemic on the three main hypothesized channels: population, human capital, and productivity.[30]

[28] This specification is similar in spirit to the standard cross-sectional growth regressions. Indeed, it benefits from and mimics a large literature that has been published on the theoretical modelling of the impact of health on economic growth. Our empirical strategy thus builds on McDonald and Roberts (2006) and Acemoglu and Johnson (2007).

[29] We present results without country fixed effects in the Appendix for comparison.

[30] Again, the estimation follows the IV set-up because of the inherent reversal causality between HIV/AIDS epidemic and human capital, productivity, and population. Therefore, our estimation frame-

So far we have said nothing about the first stage equation which is important in gauging the relevance of our instruments for causal inference. Thus, we estimate the following first-stage equation:

$$\begin{aligned} HIV_{i,j,sy} = \tau_0 + \tau_1 MC_{i,j,1959} + \tau_2 Dist_{i,j,sy} \\ + X'_{i,j,sy}\tau_3 + TE_{sy} + FE_i + \eta_{i,j,sy}, \end{aligned} \tag{6}$$

where $MC_{i,j,1959}$ is the regional shares of pre-epidemic male circumcision practices for countries in our sample. $Dist_{i,j,sy}$ is the travel time in days between region i in country j to Kinshasa during the survey year sy. As above, $X_{i,j,sy}$ is a column vector with the same controls and including initial income per capita. The rest of the variables are the same as above. $\eta_{i,j,sy}$ captures other unobservable factors affecting HIV prevalence rates in region i across countries j over the survey years sy. The underlying hypotheses for our parameters of interest are $\tau_1 < 0$ and $\tau_2 < 0$.[31] Now that we have defined our empirical model, we proceed to describe the instrumental variables.

5 Instrumental Variables

A credible IV strategy in our context requires instrument(s) that are correlated with HIV prevalence but not with unobserved factors that affect economic growth. As previously noted, our identification strategy uses two instruments.[32] First, we exploit the variation in pre-epidemic male circumcision rates across regions. Second, we use the exogenous variation in distance to the first HIV epidemic outbreak. We now present more details on the choice and the relevance of our instruments.

work takes the following form.

$$\begin{aligned} \Gamma_{i,j,sy,2012} = \alpha_0 + \alpha_1 Lightpc_{i,j,sy} + \alpha_2 \hat{HIV}_{i,j,sy} \\ + X'_{i,j,sy}\alpha_3 + TE_{sy} + FE_i + \varepsilon_{i,j,sy}, \end{aligned} \tag{5}$$

where, Γ represents output growth, population growth, human capital, and output per worker growth (our measure of productivity) with α_2 being the parameter of interest.TE_{sy} and FE_i are time and country fixed effects, respectively. ε is an error term.

[31] A standard practice in cross-sectional analysis is to cluster standard errors to adjust for intra-cluster correlations. Since our analytical framework is at the regional level with only one observations per survey year, meaningful clustering should be done at the country level. However, since we only have 25 countries, our analysis is limited by few clusters. In such cases Cameron and Miller (2010) advice against clustering as it is likely to cause even more problematic coefficients inference based on the standard errors. Therefore, the precision of our coefficient estimates strictly relies on White-Huber heteroskedastic robust standard errors, cautiously noting that the reported standard errors may be slightly inflated in the absence of clustering.

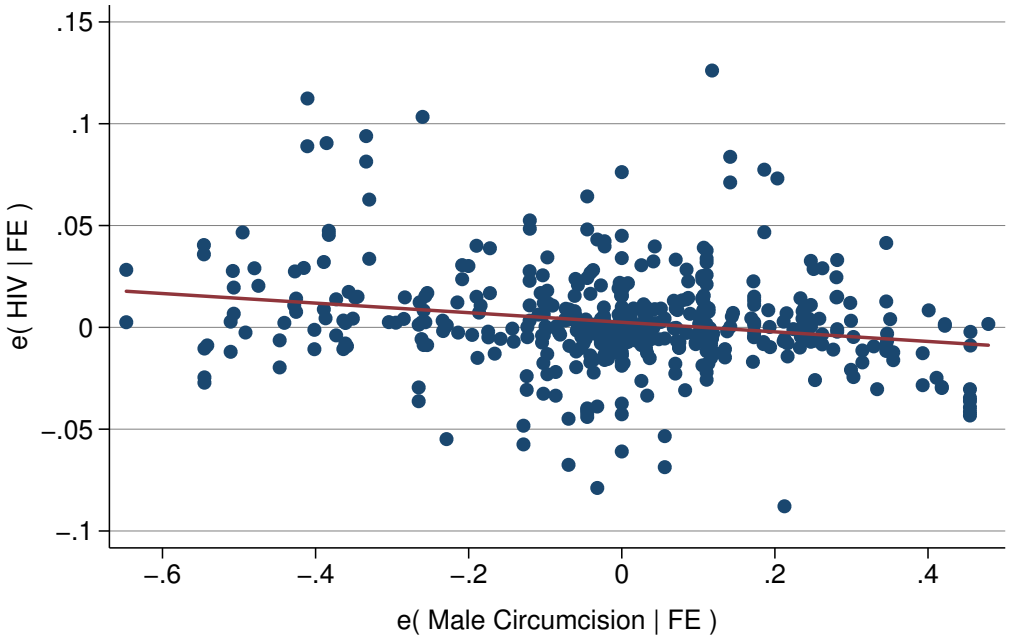
[32] We do not exploit sexual behaviour as a potential instrument as Oster (2012) because of the lack of consistent predetermined data. However, we try to include a proxy of sexual behaviour (measured as the average number of sexual partners excluding the spouse) as a third potential instrument. The first stage results becomes marginally stronger and the second stage results are similar. See Tables 19 and 20 in the appendices.

5.1 Male Circumcision

As mentioned above, medical research has demonstrated that HIV rates are lower in populations that traditionally practice male circumcision than in populations in which the procedure is rare (see, for example, Timberg and Halperin, 2012; Epstein, 2007; Donoval et al., 2006). Specifically, Donoval et al. (2006, p.386) report that for uncircumcised males "the inner mucosal[33] surface of the penis foreskin contains cells that make it highly susceptible to HIV infection."

More recently, the causal relationship between male circumcision and HIV has been established through three separate randomized controlled trials showing that the risk of infection is up to 60 percent lower for circumcised than uncircumcised men (Potts et al., 2008; Werker et al., 2006). These results have led WHO and UNAIDS to promote clinical male circumcision in similar spirit to the promotion of condom use and behaviour change.[34] In its most recent report on the global HIV epidemic, UNAIDS (2013) state that "It is projected that circumcising 80 percent of all uncircumcised adult men in the countries with low prevalence of male circumcision by 2015 would avert one in five new HIV infections by 2025, with long-term prevention benefits for women as well as men."

Figure 3: Partial Correlation: HIV Prevalence and Male Circumcision



Source: Authors' calculations

[33] Also known as mucous membrane which refers to the head of the penis and the inner layer of the foreskin.

[34] More details available at: `www.who.int/hiv/topics/malecircumcision/en/`

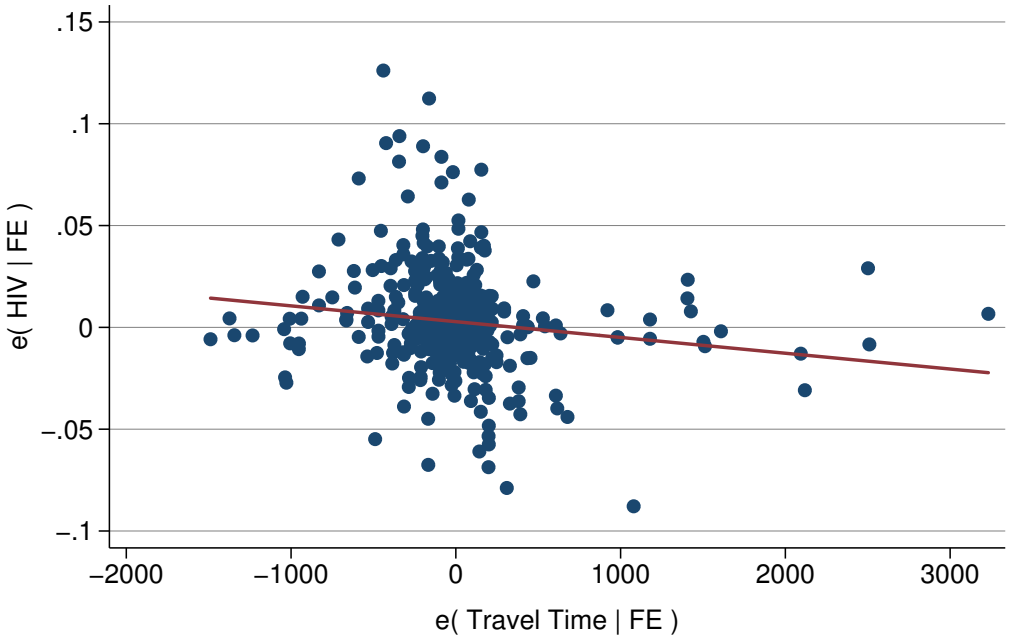
The choice of male circumcision as an instrument builds on these scientific arguments. This choice is also warranted by our data. Figure 3 plots the partial correlation between our pre-epidemic male circumcision measure and the HIV prevalence epidemic net of country fixed effects. The correlation coefficient is -0.023 and strongly significant.

With exception of Swaziland and Burundi in the upper and lower bounds , respectively, a general picture portrayed by Table 2 is that in the period of our analysis, male circumcision is low in countries with high epidemic prevalence relative to those with low prevalence rates. Again, this signals the relevance of male circumcision as our instrument.

5.2 Distance Measures

The underlying assumption for the choice of distance as an instrument is also related to the spread of infectious diseases. In this paper we aim to realistically capture the HIV/AIDS spread from the origin to each region across countries. We therefore estimate average travel times in days from each region to Kinshasa in DRC. To better trace the spread, we also account for differences in population densities in our calculations of travel times in days – c.f. the appendices for more details of the estimation procedures.

Figure 4: Partial Correlation: HIV Prevalence and Travel Time (in days)



coeff=0.022, se=0.0067, t=3.27

Source: Authors' calculations

The idea of using distance as an instrument is not new. As mentioned earlier on, Oster (2012) proposes and uses the great circle distance from the origin in the center of DRC to the centroid of other countries as an instrument for HIV prevalence rates to estimate the causal impact of HIV prevalence on sexual behaviour across a sample of 14 countries in sub-Saharan Africa. As mentioned in Section 3, we use the city of Kinshasa as our main epicentre since most recent evidence points in that direction (see Timberg and Halperin, 2012, for details). We then use Oster (2012)'s measure for robustness checks. All the results using the alternative distance measures are shown in the appendices.

Figure 4 shows the partial correlation between our distance measure and HIV prevalence rates net of country specific effects. The coefficient is -0.022, strong and significant. The last column in Table 2 confirms the same pattern although the relation is not clear at the national level. Overall, in correlation terms, distance qualifies as a relevant instrument for asserting causal association between HIV/AIDS and economic growth.

Table 2: HIV Prevalence rates, Male circumcision and Weighted Distance

Country	Year	HIV Prevalence(%)			Instruments	
		Total	Female	Male	MC (%)	Travel Time (days)
Swaziland	2006	25.9	31.1	19.7	96.8	1.74
Lesotho	2009	23.0	26.7	18.0	0.0	1.90
Zimbabwe	2010	15.2	17.7	12.3	39.7	1.44
Zambia	2007	14.3	16.1	12.3	33.3	1.23
Mozambique	2009	11.4	13.0	9.2	52.8	2.05
Malawi	2010	10.6	12.9	8.1	24.8	1.90
Uganda	2011	7.3	8.3	6.1	9.3	1.26
Kenya	2008	6.3	8.0	4.3	63.1	1.58
Tanzania	2007	5.7	6.6	4.6	44.3	1.89
Cameroon	2011	4.3	5.6	2.9	72.7	0.85
Cote d'Ivoire	2005	4.6	6.2	2.8	57.0	2.13
Congo	2009	3.2	4.1	2.1	79.4	0.51
Rwanda	2010	3.0	3.7	2.2	2.5	1.25
Ghana	2003	2.2	2.7	1.5	38.0	1.85
Liberia	2007	1.6	1.9	1.2	75.6	2.27
Sierra Leone	2008	1.5	1.7	1.2	89.4	2.41
Guinea	2005	1.5	1.9	0.9	79.5	2.33
Ethiopia	2011	1.5	1.9	1.0	69.5	2.14
Burundi	2010	1.4	1.7	1.0	2.1	1.24
Mali	2006	1.3	1.5	1.0	87.0	2.11
Congo DRC	2007	1.3	1.6	0.9	67.2	0.79
Benin	2006	1.2	1.5	0.9	62.1	1.45
Burkina Faso	2010	1.0	1.2	0.8	76.8	1.81
Niger	2006	0.7	0.7	0.8	85.7	1.61
Senegal	2010	0.7	0.8	0.5	81.8	2.49
Total		5.9	7.1	4.6	55.7	1.67

To derive consistent estimates for causal inference of the epidemic on economic growth our instruments must satisfy two conditions: *relevance* and the *exclusion restriction*. The relevance condition requires correlation between the instruments and the endogenous variable. We have tested and shown the relevance of our instruments. In

the results section we will formally test the strength of this restriction. The exclusion restriction requires that instruments are not correlated with unobserved characteristics that are orthogonal to our outcome variable, i.e., output per capita growth. We cannot test this restriction, but we examine potential factors that could violate it. Thus, we describe the potential threats to our identification strategy in the following subsection.

5.3 Identification Threats

In relation to male circumcision, several confounding factors are likely to threaten the exclusion restriction. The main factors are religion and ethnicity.[35] Religion becomes a potential threat when male circumcision is performed as a religious ritual – which is particularly the case in Muslim populations (Werker et al., 2006) – and if religion at the same time has an impact on economic activity in itself (see, for example, Barro and McCleary, 2003). By no surprise, our data show that regions with higher male circumcision ratios have higher shares of Muslims. To address this potential identification threat we control for a region's share of Muslim population in our model specifications.

Similarly, ethnicity becomes a threat to identification if specific tribes, by tradition, practice circumcision as part of their cultural identity and this has an independent impact on economic outcomes in a similar fashion as religion. Michalopoulos and Papaioannou (2013) find a significant association between ethnic political centralization and contemporary regional economic growth (also measured by the growth of light intensity). They use measures of ethnicity in pre-colonial Africa based also on data from Murdock et al. (1959). As noted before, the measure is an index ranging from 0 to 4 with 0 being the lowest level with no political autonomy beyond community and 4 being the highest level which indicates membership of a large state. We use Michalopoulos and Papaioannou (2013)'s ethnicity measure to thwart the identification threats in our estimations: the measure allows controlling for otherwise unobserved confounding bias that would likely threat our identification strategy. One may naturally argue that the spatial distribution of African ethnicities has changed since Murdock et al. (1959) collected their data. However, Nunn (2008) shows a relatively strong and significant correlation of 0.55 between the historic and present day ethnicities measured using Afro barometer data and Murdock et al. (1959)'s data.

As with circumcision, distance is also prone to identification threats. Normally it is difficult to formally test the exclusion restriction but, in economic terms, we can investigate whether regions far from Kinshasa are significantly different from regions that are in close proximity. We could test for potentially omitted factors that affect

[35] A potential third factor is conflicts. We have investigated conflicts and find no relationship between our instruments and regional conflicts. If anything, we find a negative correlation consistent with Potts et al. (2008). However this relationship is mainly driven by long-lasting conflicts in low HIV prevalence countries like Ethiopia and Burundi.

growth and at the same time are correlated with distance. Our tests (unreported) give us no reason to believe that regions far from Kinshasa are significantly different from regions that are close.[36]

6 Results

6.1 Male Circumcision, Distance and the HIV/AIDS Epidemic

In this section we show the extent to which our instruments are predictors of the HIV epidemic. Table 3 presents the first stage estimates.

Table 3: First stage: Male Circumcision, Distance, and HIV
Dependent Variable: HIV Prevalence

	Model 1	Model 2
Initial Income	0.220**	0.062
	[0.106]	[0.112]
Male Circumcision	-0.023***	-0.020**
	[0.008]	[0.009]
Travel Time (in days)	-0.020***	-0.010*
	[0.007]	[0.006]
Muslim Shares		-0.026***
		[0.009]
Capital Region Indicator		0.013**
		[0.006]
Ethnicity Controls	No	Yes
Time Fixed Effects	Yes	Yes
Country Fixed Effects	Yes	Yes
Regions	241	236
F-Statistic	68.897	70.003
R-squared	0.870	0.883

Notes: This table shows the first stage regression estimates of male circumcision and distance to the first outbreak on HIV prevalence across 25 countries in sub-Saharan Africa for the period 2003 – 2012. Controlling for initial income, Muslim shares and ethnicity levels, overall the table establishes a statistically significant negative effect of male circumcision and distance on on HIV. Standard errors in brackets, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Column 1 shows the regression estimates conditional on initial income, time and country fixed effects. The estimates show that male circumcision and distance are strong and negatively related to HIV prevalence. Column 2 shows the estimates conditional on covariates in addition to initial income, time fixed effects and country fixed effects.

[36] Results can be obtained from the authors upon request.

Both magnitude and precision decline slightly but remain significant at 5 percent level for male circumcision and marginally significant for distance. Overall, the results show that the effect of our instruments on HIV prevalence rates is negative and consistent with the conjectured underlying hypothesis.

We now turn our attention to the effect of other covariates in the first stage regressions. Except for Muslim share and capital region indicator which are statistically significant, the rest of the covariates have insignificant effects on HIV prevalence rates. [37] The sign of the coefficient on Muslim share is negative and therefore consistent with our underlying conjecture in the way it affects HIV prevalence through male circumcision practised by tradition. The indicator for capital region is positive, suggesting that HIV is more prevalent in capital regions and urban areas relative to other regions that are further away from the capitals.[38]

6.2 HIV/AIDS Impact on Economic Growth

This section reports the estimates showing the causal effect of the HIV/AIDS epidemic on income per capita growth (i.e., second stage regression estimates). The estimates show that the sign on the HIV prevalence coefficient is negative but statistically insignificant. Controlling for time and country fixed effects, this cross-sectional evidence suggests that the epidemic has had insignificant effect on income per capita growth in 241[39] regions across 25 countries in sub-Saharan Africa during 2003–2012. However, if the analysis is not conditioned on country fixed effects, we find marginal effects of the epidemic on economic growth in the region during the same period – c.f. Column 4 of Table 10 in the appendices. We argue that this marginal effect potentially explains the negative effect that has been documented in some previous cross-country studies. Note one caveat with our estimates: they are cross-sectional and do not allow causal inferences of the epidemic and economic growth within regions across countries.

In Table 4, Columns 1 and 2 report, respectively, OLS and 2SLS results without other covariates. As noted, the estimates suggest a sizeable but insignificant effect of the epidemic on growth per capita. However, alluding to Section 5.3 we introduce the relevant controls in Columns 3 and 4. Even after the inclusion of these controls, the results in these Columns (relative to Columns 1 and 2) still report high, negative, but statistically insignificant effects of the epidemic on income per capita growth. Note, however, while all regressions pass tests for both endogenous regressors and over-identification, the Cragg-Donald or Kleibergen-Paap's F statistic of the first stage estimation (which is

[37] Except for the most politically centralized ethnic groups, however, this result should be interpreted with high caution as only one region within our sample (in Ethiopia) belongs to this ethnic category.

[38] We also tried controlling for distance to capital instead of the indicator for being in the country's capital region. Distance to capital was highly statistically negative. The results did not change.

[39] We lose 5 regions when including all covariates. This is caused by lack of Muslim share data in these five regions.

less than the rule of thumb of 10) casts some doubt on the relevance of our instruments. Following a suggestion by Cameron and Trivedi (2005), if we impose a less conservative rule of thumb (Cragg-Donald or Kleibergen-Paap's F statistic > 5) which allows the IV bias to be 20 percent that of OLS, then we can argue for the relevance of our instruments (in Table 4, all the F-statistics are > 5).

Table 4: HIV Prevalence and Output per Capita Growth
Dependent Variable: Output Per Capita Growth

	OLS	2SLS	OLS	2SLS	LIML	GMM
HIV Prevalence	-0.347	-1.167	-0.490*	-1.945	-2.029	-2.062
	[0.259]	[0.945]	[0.286]	[1.682]	[1.784]	[1.677]
Initial Income	-2.971***	-2.711***	-2.746***	-2.569***	-2.559***	-2.662***
	[0.911]	[0.884]	[0.957]	[0.916]	[0.920]	[0.909]
Muslim Share			-0.034	-0.084	-0.087	-0.090
			[0.047]	[0.077]	[0.080]	[0.077]
Capital Region Indicator			-0.031	-0.012	-0.011	-0.010
			[0.019]	[0.028]	[0.029]	[0.028]
Ethnicity Controls	No	No	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Regions	241	241	236	236	236	236
F-Statistic	11.118	9.064	8.425	7.618	7.607	7.646
R-squared	0.435	0.419	0.465	0.423	0.418	0.415
Cragg-Donald or Kleibergen-Paap		6.565		4.498	4.498	4.498
p-value Anderson-Rubin LR					0.484	
p-value Anderson-Rubin F-test		0.405		0.385	0.385	0.385
p-value Hansen J statistic		0.472		0.402	0.404	0.402

Notes: This table shows the second stage regression estimates of HIV prevalence – instrumented by male circumcision and distance to the first epidemic outbreak – on economic growth across 25 countries in sub-Saharan Africa for the period 2003-2012. Economic growth indicator is output per capita growth proxied by average growth in night lights per capita. Controlling for initial income, Muslim shares and ethnicity levels, overall, the table establishes an insignificant negative effect of the HIV prevalence rates on the growth of output per capita. The Cragg-Donald or Kleibergen-Paap is an F statistic for weak identification under a null of a weakly identified equation. The Anderson-Rubin LR is an over-identification test for weak instrument using the LIML estimator. The Anderson-Rubin F-test is test of the significance of endogenous regressors under the null that endogenous regressors in the structural equation are jointly equal to zero. The Hansen J statistic tests the null that instruments are uncorrelated with the errors. Standard errors in brackets, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

To ascertain the relevance of our instruments, we follow Greene (2012) and fit the limited information maximum likelihood estimator (LIML) which has asymptotic properties similar to 2SLS but is more robust in the presence of weak instruments. LIML has also been proven more robust when the sample size is relatively small. For comparison of estimates Greene (2012) also recommends fitting a GMM specification which is more robust in estimating the covariance variance matrix compared to both 2SLS and LIML approaches. Column 5 reports the results based on LIML and Column 6 reports results based on GMM. The coefficient estimates of 2SLS, LIML, and GMM are indeed

closely the same.[40] Taken together, both LIML and GMM specifications confirm the 2SLS findings and leads us to confirm our earlier assertion that HIV/AIDS epidemic has had statistically insignificant effect on income per capita growth in sub-Saharan Africa between 2003 and 2012.

To test the relevance of our instruments further we also include an additional test by following Angrist and Pischke (2008) who, in the presence of weak instruments, suggest the inclusion of the instruments separately. Tables 15 and 16 in the appendices report the findings based on this estimation strategy. Compared to including both instruments at once, the first stage F-statistic increases if we use only male circumcision and declines using only distance. However, including just one instrument does not change the main result: negative but insignificant effects of HIV/AIDS on economic growth.

Turning to the effects of other covariates in the models, Table 4 shows that initial income negatively (in convergence sense) affects economic growth.[41] An interesting observation, moreover, is that coefficient estimates on Muslim shares, ethnicity measures, and the indicator for capital region, which affect economic growth through the HIV epidemic, show no significant effect on per capita output growth. Their inclusion in the analysis, however, corrects for potential endogenous instruments and over-identification restrictions.

6.3 Mechanisms

The coefficients from our main results potentially mask a myriad of policy relevant information. We now present the results on total income growth as well as the hypothesized mechanisms through which HIV affects economic outcomes – namely human capital, productivity, and population growth. Although our baseline estimates suggest insignificant effects of HIV on growth, we investigate these mechanisms to see if we find a similar result.

Table 5 shows the estimation for total income growth. The table shows similar results to the one on income per capita growth: an insignificant effect of the HIV epidemic on economic growth. Also, the controls have similar insignificant effect but serve as controls in the first stage to account for unobserved confounding bias.

Human capital has a direct link with economic growth outcomes. Naturally, the extent to which it is affected by the HIV epidemic can also be reflected in its effects on economic growth outcomes. Columns 1 - 4 in Table 6 present our regression estimates on human capital (proxied with average years of education). Our main finding is that the epidemic appears to have had a negative, but statistically insignificant, impact on

[40] This is also the case for all of the following regressions and hence we do not report the LIML and GMM estimates in the remaining tables.

[41] Since the effect of initial income looks rather large, we tried to exclude it as a check. The results remain qualitatively unchanged.

human capital levels/[42] This is generally true in all our model specifications. Inferences on the covariates show that initial income positively affects human capital outcomes. This result is not surprising: higher levels of initial endowments tend to go hand in hand with higher levels of education.[43] The table also reports three other observations. First, Muslim shares which, as stated before, might affect growth through the HIV epidemic, appear to have a robust negative effect on human capital accumulation. What precise mechanisms these estimates suggest is not clear. Second, human capital outcomes tend to improve in regions that are capital regions relative to those that are not. This is expected: the incentives, as well as opportunities for acquiring human capital in cities are greater than in other regions as the payoffs in cities are likely to be bigger.

Table 5: Total HIV Prevalence and Output Growth
Dependent Variable: Output Growth

	OLS	2SLS	OLS	2SLS	LIML	GMM
HIV Prevalence	-0.380	-1.336	-0.519*	-2.209	-2.297	-2.299
	[0.243]	[0.930]	[0.276]	[1.649]	[1.741]	[1.645]
Initial Income	-3.075***	-2.771***	-2.918***	-2.713***	-2.703***	-2.789***
	[0.890]	[0.869]	[0.934]	[0.896]	[0.900]	[0.891]
Muslim Share			-0.027	-0.085	-0.088	-0.091
			[0.045]	[0.074]	[0.076]	[0.073]
Capital Region Indicator			-0.024	-0.003	-0.001	-0.001
			[0.019]	[0.027]	[0.028]	[0.027]
Ethnicity Controls	No	No	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Regions	241	241	236	236	236	236
F-Statistic	11.292	8.497	8.371	6.409	6.337	6.457
R-squared	0.447	0.425	0.474	0.415	0.409	0.408
Cragg-Donald or Kleibergen-Paap		6.565		4.498	4.498	4.498
p-value Anderson-Rubin LR					0.504	
p-value Anderson-Rubin F-test		0.329		0.326	0.326	0.326
p-value Hansen J statistic		0.478		0.426	0.427	0.426

Notes: This table shows the second stage regression estimates of HIV prevalence – instrumented by male circumcision and distance to the first epidemic outbreak – on economic growth across 25 countries in sub-Saharan Africa for the period 2003 – 2012. Economic growth indicator is output growth proxied by average growth in night light intensity. Controlling for initial income, Muslim shares and ethnicity levels, overall, the table establishes an insignificant negative effect of the HIV prevalence rates on the growth of output. The Cragg-Donald or Kleibergen-Paap is an F statistic for weak identification under a null of a weakly identified equation. The Anderson-Rubin LR is an over-identification test for weak instrument using the LIML estimator. The Anderson-Rubin F-test is test of the significance of endogenous regressors under the null that endogenous regressors in the structural equation are jointly equal to zero. The Hansen J statistic tests the null that instruments are uncorrelated with the errors. Standard errors in brackets, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

[42] When estimating the model using simple OLS, the estimates are positive and significant. Thus, this can potentially lead to wrong conclusions because of the OLS bias.

[43] Note that we cannot control for initial educational levels in this set-up due to data limitations. We would however expect initial income, at least to some degree, to capture some of the variation.

Table 6: Total HIV Prevalence and Human Capital, Productivity, and Population growth
Dependent Variables: Human Capital (i.e.,Education), Output Per Worker (i.e., Productivity) and Population Growth

	Human Capital				Productivity				Population Growth			
	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS
HIV Prevalence	12.947***	3.959	6.925**	-30.948	0.068***	0.088*	0.063***	0.096	-0.033	-0.169*	-0.029	-0.264
	[3.623]	[9.894]	[3.092]	[20.164]	[0.019]	[0.053]	[0.019]	[0.079]	[0.063]	[0.097]	[0.051]	[0.237]
Initial Income	85.079***	87.934***	49.551***	54.150***	0.118**	0.111*	0.011	0.007	-0.103	-0.060	-0.172	-0.144
	[17.245]	[15.976]	[12.923]	[12.749]	[0.059]	[0.059]	[0.048]	[0.046]	[0.133]	[0.115]	[0.140]	[0.117]
Muslim Share			-2.329***	-3.640***			0.001	0.002			0.007	-0.001
			[0.498]	[1.027]			[0.001]	[0.003]			[0.015]	[0.020]
Capital Region Indicator			2.114***	2.602***			0.006***	0.006***			0.007	0.010*
			[0.301]	[0.411]			[0.001]	[0.001]			[0.004]	[0.006]
Ethnicity Controls	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regions	241	241	236	236	241	241	236	236	241	241	236	236
F-Statistic	48.632	51.373	76.077	42.283	7.925	8.312	8.820	8.999	162.796	81.976	53.728	41.491
R-squared	0.780	0.772	0.860	0.745	0.583	0.581	0.622	0.616	0.491	0.477	0.506	0.469
Cragg-Donald or Kleibergen-Paap		6.565		4.498		6.565		4.498		6.565		4.498
p-value Anderson-Rubin F-test		0.044		0.149		0.200		0.525		0.221		0.578
p-value Hansen J statistic		0.010		0.987		0.575		0.707		0.903		0.713

Notes: This table shows the second stage regression estimates of HIV prevalence – instrumented by male circumcision and distance to the first epidemic outbreak – on the following potential channels: columns 1 – 4 average years of education (human capital), Columns 5 – 8 output per worker (productivity), Columns 9 – 12 population growth. Controlling for initial income, Muslim shares and ethnicity levels, overall, the table establishes an insignificant negative effect of the HIV prevalence rates on the three potential mechanisms. Standard errors in brackets, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Similar to human capital, labour productivity also has a direct causal link with economic growth outcomes to the extent that if affected by the HIV/AIDS epidemic the effect can also be passed on to economic outcomes. Columns 5 - 8 report regressions based on output per worker – our measure of labour productivity. The results reveal very small positive – but when controlling for all covariates – insignificant effects. Although statistically weak, the positive signs on the estimates are consistent with Young (2005). The table also reports estimates on other covariates. While the effect of initial income disappears with the introduction of covariates in Columns 7 - 8, no effect is revealed on Muslim share and measures of ethnicities. However, the coefficient on the capital region indicator is positive, robust, and significant. We interpret this finding in the same spirit as we did for human capital: labour productivity is greater in capital regions relative to other regions.

Finally, Columns 9 - 12 show regression estimates for population growth. One could imagine two opposing effects of HIV/AIDS on population growth: HIV/AIDS could, through the increase in mortality, lead to a lower population growth. However, HIV prevalence could also lead to a higher population growth, for example through a precautionary demand for children leading to higher fertility (Kalemli-Ozcan, 2012), or because of increased access to antiretroviral treatment, that lowers mortality. Our findings show a negative but statistically insignificant effect of HIV prevalence on population growth.

Overall, all the regression results presented in this section are robust to all specification checks. The findings show support to the claim that the HIV/AIDS epidemic effects on total growth, human capital, labour productivity, and population growth are insignificant. The is corroborate our main finding that the epidemic has not been a threat to economic growth in sub-Saharan Africa in recent years.

6.4 HIV/AIDS Impact on Regional Inequality

Although we find insignificant effects on economic growth, human capital, productivity, and population growth of the epidemic it still remains to rule out similar effects on poverty and inequality. For instance, Barro (2000) shows that the epidemic increased inequality. Similarly, Werker et al. (2006) show weak indications of increased poverty (measured by malnutrition). Due to limitations on regional level poverty data, we restrict our analysis to regional inequality. As mentioned before, we exploit the spatial nature of the light data to construct lights intensity spatial Gini index which we use a proxy for regional income differences.

Table 7: Effect on Regional Inequality
Dependent Variable: Changes in Inequality measured by Spatial Lights Gini Index

	OLS	2SLS	OLS	2SLS	LIML	GMM
HIV Prevalence	-0.065	-0.471	-0.036	-0.155	-0.194	-0.132
	[0.087]	[0.347]	[0.074]	[0.482]	[0.635]	[0.482]
Initial Income	-0.652	-0.523	-0.879	-0.865	-0.860	-0.435
	[1.139]	[1.036]	[1.261]	[1.154]	[1.150]	[1.121]
Muslim Shares			0.039**	0.035	0.033	0.038
			[0.018]	[0.024]	[0.029]	[0.024]
Capital Region Indicator			0.010	0.012	0.012	-0.000
			[0.023]	[0.022]	[0.023]	[0.021]
Ethnicity Controls	No	No	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Regions	241	241	236	236	236	236
F-Statistic	2.836	1.796	3.000	3.145	3.283	3.088
R-squared	0.169	0.117	0.226	0.223	0.220	0.207
Cragg-Donald or Kleibergen-Paap		6.565		4.498	4.498	4.498
p-value Anderson-Rubin F-test		0.185		0.315	0.315	0.315

Notes: This table shows the second stage regression estimates of HIV prevalence – instrumented by male circumcision and distance to the first epidemic outbreak – on the change in the Gini coefficient across 25 countries in sub-Saharan Africa for the period 2003 – 2012. The dependent variable is calculated as the average yearly change in Gini. Controlling for initial income, Muslim shares and ethnicity levels, overall, the table establishes a negative but insignificant effect of the HIV prevalence rates on the Gini coefficient. Standard errors in brackets, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

We apply the same analytical framework but with changes in spatial Gini[44] as an outcome variable. The changes are calculated as the average annual changes in spatial Gini coefficient between the DHS year and 2012. Table 7 presents our findings. The estimates, though statistically insignificant, suggest that a percentage point increase in the epidemic prevalence rates generally lowered spatial inequality between regions across all model specifications. We speculate that the onset and the persistence of the epidemic in Africa, which as previously noted affected mainly the most productive age group, reduced inequality across regions by levelling income differences across and between regions. Our findings somewhat echo those of Werker et al. (2006).

7 Robustness Checks

We subject our empirical model to various sensitivity tests in order to verify our main regression results. First, we run regressions without country fixed effects to study the importance of accounting for these and potentially what conclusions cross-country analysis

[44]Calculated as shown in equation 5.

would find. We find negative effects of HIV/AIDS on economic growth (see Column 4 of Table 10), but the result is not robust.[45] Moreover, the result shows the importance of controlling for country-specific effects.

Next, we fit the baseline model using Oster (2012)'s distance origin. The results are very similar. The first and second stage estimates are presented in Table 11 and 12, respectively. These results are robust to a number of checks such as the great circle distance and population weighed and unweighed measures.[46]

In Table 2 we showed that females' HIV prevalence rates are significantly higher than males' in virtually all countries. We thus evaluate whether there have been differential growth effects of female and males HIV prevalence rates. Table 13 presents the respective first stage results. Note that the effect of *male* circumcision is stronger for women. Our finding is consistent with Haacker (2004) who argue that *male-to-female* transmission rates in sub-Saharan Africa are twice the *female-to-male* transmission rates. The second stage results are presented in Table 14 in the appendices. Columns 1–4 present the estimates for female HIV prevalence rates. Columns 5–8 present the estimates for male HIV prevalence rates. The biased OLS estimates show negative and significant coefficients. Correcting for this bias, the IV estimates still show negative but insignificant results. Finally, Tables 17 and 18 show the results when we pool our samples in countries with multiple surveys. The results remain unchanged.

In general, all our results point to the same conclusions. First, we find statistically insignificant effects of the epidemic on economic growth and regional income disparities in sub-Saharan Africa between 2003 and 2012. Second, we find male circumcision a strong predictor of the epidemic during this period. And third, distance is a weak predictor, arguably because of the epidemic's maturity during this period.

8 Conclusions

This paper revisits and investigates the debate on the impact of the HIV/AIDS epidemic on economic growth in sub-Saharan Africa. Our causal identification strategy instruments the HIV/AIDS epidemic with predetermined male circumcision rates and distance to first epidemic outbreak. This IV strategy mitigates endogeneity concerns between health and economic growth.

While the use of these instruments is not new, we contribute to the existing literature in three important ways. First, we combine our instruments together, which to the best of our knowledge, has never been done before. Second, our level of analysis

[45] When we estimate our model without country fixed effects the sign on distance in the first stage becomes positive (see Table 9). This is not in line with the proposed mechanism. However if we estimate the same model using only male circumcision results are very similar – i.e., HIV is still borderline significant in the second stage.

[46] The specification check results are not reported but available upon request from authors.

is sub-national (i.e., regions) unlike the previous studies that used countries as unit of analysis. The focus on sub-national analysis allows us not only estimate the effects at regional level but also control for unobserved country-specific effects potentially affecting both HIV/AIDS and economic growth which is impossible for country-level analysis. Third, our identification strategy accounts for measures of ethnicity and religion, which are potential threats to identification, rendering inference more close to the real world reality.

The results show that male circumcision is a strong and robust predictor of HIV prevalence in our sample. This finding resonates well with the ongoing male circumcision programs led by WHO and UNAIDS in sub-Saharan Africa. However, our distance measure adds less explanatory power. This finding is unsurprising, especially when including country fixed effects (which controls for unobserved country-specific heterogeneity) in the estimated model. Moreover, we attribute distance's lower explanatory power to the maturity of the epidemic since its initial outbreak.

Our main finding is that the effect of the HIV/AIDS epidemic on economic growth outcomes is negative but statistically insignificant. The same effect holds when checked against human capital, output per worker and population growth which are potential pathways for the epidemic to affect economic growth. One possible explanation for this insignificant effect – though we cannot verify it with our data – is the increasing availability of antiretroviral therapy (ART), which prolongs lives and thus potentially keeps people productive. Since 2005, massive investments have been made in scaling up access to ARTs and the average coverage rate for adults reached 68 percent in 2012 (UNAIDS, 2013). Our data overlaps with this period of massive ARTs investments thus supporting our plausible belief on ARTs role in diminishing the epidemic's negative economic consequences.

Does this mean HIV/AIDS does not matter for economic development in sub-Saharan Africa? We cannot completely rule out that HIV/AIDS is economically insignificant. Importantly, there could still be multiple ways, especially at the micro level, in which the epidemic can negatively affect social and economic indicators. By nature of our empirical design, we cannot rule out such potential micro-level effects.

Our results are in line with other studies that examine the impact of the HIV epidemic on economic growth and find little or no effect. For example, Werker et al. (2006) find insignificant effects of the epidemic on economic growth for a sample of 32 African countries. Moreover, this paper contributes to a broad debate on disease and development, particularly in developing countries. Specifically, the findings are in line with recent studies, for example Acemoglu and Johnson (2007) and Soares (2007), that also find little or no effect of health outcomes on economic indicators in the developing world[47]. Broadly, this paper relates to a burgeoning literature on deep factors under-

[47] Empirical evidence on the developed world also shows similar conclusions – Hansen (2014) presents

lying cross-country income differences around the globe for which diseases have been attributed to hold substantial explanatory power – (see, for example Weil, 2010; Bloom et al., 2009; Lorentzen et al., 2008, among others).

In summary, this paper revisited and investigated the lingering debate on HIV/AIDS and economic growth in 241 regions across 25 countries in sub-Saharan Africa during 2003–2012. We argued that the HIV/AIDS epidemic did not affect economic growth in the region during this period, perhaps because of the increased uptake of ARTs. Nonetheless, we still believe HIV/AIDS has the potential to affect social and economic outcomes in the continent. Going forward we believe future research can be of interest in evaluating the extent to which antiretroviral therapy (ART) has abated the effects of the epidemic on social and economic outcomes.

References

Acemoglu, D. and S. Johnson (2007). Disease and development: The effect of life expectancy on economic growth. *Journal of Political Economy vol. 115, no. 6*.

Ajakaiye, O. and M. Audibert (2010). Endemic Diseases and Development: Introduction and Overview. *Journal of African Economies 19*(suppl 3), 3–11.

Alemu, Z. G., T. L. Roe, and R. B. Smith (2005). The Impact of HIV on Total Factor Productivity. *Working paper*. Working paper.

Alesina, A. F., S. Michalopoulos, and E. Papaioannou (2015). Ethnic inequality. *Journal of Political Economy 123*(3), 547–724.

Angrist, J. D. and J.-S. Pischke (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Barro, R. J. (2000). Inequality and Growth in a Panel of Countries. *Journal of economic growth 5*(1), 5–32.

Barro, R. J. and R. M. McCleary (2003). Religion and Economic Growth across Countries. *American Sociological Review 68*(5), pp. 760–781.

Bleakley, H. (2010). Health, human capital, and development. *Annu. Rev. Econ. 2*(1), 283–310.

Bloom, D. E., D. Canning, and G. Fink (2009). Disease and development revisited. Technical report, National Bureau of Economic Research.

recent work in this regard.

Bloom, D. E. and A. S. Mahal (1997). Does the AIDS epidemic threaten economic growth? *Journal of Econometrics 77*(1), 105–124.

Bonnel, R. (2000). HIV/AIDS and Economic Growth: A Global Perspective. *South African Journal of Economics 68*(5), 360–379.

Cameron, A. C. and D. L. Miller (2010). Robust inference with clustered data. *Handbook of empirical economics and finance*, 1–28.

Cameron, A. C. and P. K. Trivedi (2005). *Microeconometrics: methods and applications*. Cambridge university press.

Chen, X. and W. D. Nordhaus (2011). Using luminosity data as a proxy for economic statistics. *Proceedings of the National Academy of Sciences 108*(21), 8589–8594.

Cuddington, J. T. (1993a). Further results on the macroeconomic effects of AIDS: the dualistic, labor-surplus economy. *The World Bank Economic Review 7*(3), 403–417.

Cuddington, J. T. (1993b). Modeling the Macroeconomic Effects of AIDS, with an Application to Tanzania. *The World Bank Economic Review 7*(2), 173–189.

Dixon, S., S. McDonald, and J. Roberts (2002). The impact of HIV and AIDS on Africa's economic development. *BMJ: British Medical Journal 324*(7331), 232.

Donoval, B. A., A. L. Landay, S. Moses, K. Agot, J. Ndinya-Achola, E. A. Nyagaya, I. MacLean, and R. C. Bailey (2006, Mar). HIV-1 Target Cells in Foreskins of African Men With Varying Histories of Sexually Transmitted Infections. *American Journal of Clinical Pathology 125*(3), 386–391.

Epstein, H. (2007). *The invisible cure: Africa, the West, and the fight against AIDS*. Macmillan.

Greene, W. H. (2012). *Econometric analysis*. Pearson Education Limited.

Haacker, M. (2004). *Macroeconomics of Hiv/Aids*. International Monetary Fund.

Halperin, D. T. and R. C. Bailey (1999). Male circumcision and HIV infection: 10 years and counting. *Lancet 354*, 1813–15.

Hansen, C. W. (2014). Cause of death and development in the US. *Journal of Development Economics 109*(0), 143 – 153.

Henderson, J. V., A. Storeygard, and D. N. Weil (2012). Measuring Economic Growth from Outer Space. *American Economic Review 102*(2), 994–1028.

Kalemli-Ozcan, S. (2012, Jul). AIDS, reversal of the demographic transition and economic development: Evidence from africa. *Journal of Population Economics 25*(3), 871–897.

Korber, B. (2000, Jun). Timing the Ancestor of the HIV-1 Pandemic Strains. *Science 288*(5472), 1789–1796.

Lorentzen, P., J. McMillan, and R. Wacziarg (2008). Death and development. *Journal of Economic Growth 13*(2), 81–124.

McDonald, S. and J. Roberts (2006, Jun). AIDS and economic growth: A human capital approach. *Journal of Development Economics 80*(1), 228–250.

Michalopoulos, S. and E. Papaioannou (2013). Pre-Colonial Ethnic Institutions and Contemporary African Development. *Econometrica 81*(1), 113–152.

Murdock, G. P. et al. (1959). *Africa: its peoples and their culture history*. McGraw-Hill New York.

Nelson, A. (2008). Travel time to major cities: A global map of Accessibility. *Global Environment Monitoring Unit - Joint Research Centre of the European Commission, Ispra Italy*.

Nunn, N. (2008). The long-term effects of AfricaŠs slave trades. *Quarterly Journal of Economics*.

Oster, E. (2012). Hiv and sexual behavior change: Why not Africa? *Journal of Health Economics 31*(1), 35–49.

Papageorgiou, C. and P. Stoytcheva (2008). What is the impact of AIDS on cross-country income so far? Evidence from newly reported AIDS cases. Technical report, Mimeo, International Monetary Fund.

Pinkovskiy, M. and X. Sala-i Martin (2014). Lights, Camera,... Income!: Estimating Poverty Using National Accounts, Survey Means, and Lights. Technical report, National Bureau of Economic Research.

Potts, M., D. T. Halperin, D. Kirby, A. Swidler, E. Marseille, J. D. Klausner, N. Hearst, R. G. Wamai, J. G. Kahn, and J. Walsh (2008). Reassessing HIV prevention. *SCIENCE-NEW YORK THEN WASHINGTON- 320*(5877), 749.

Soares, R. R. (2007). On the determinants of mortality reductions in the developing world. *Population and Development Review 33*(2), 247–287.

Sonnet, J., J.-L. Michaux, F. Zech, J.-M. Brucher, M. De Bruyere, and G. Burtonbuy (1987). Early AIDS Cases Originating from Zaire and Burundi (1962-1976). *Scandinavian Journal of Infectious Diseases 19*, 511–517.

Timberg, C. and D. Halperin (2012). *Tinderbox: How the West Sparked the AIDS Epidemic and How the World Can Finally Overcome It*. The Penguin Press.

UNAIDS (2013). *Global report: UNAIDS report on the global AIDS epidemic: 2012*. UNAIDS.

Weil, D. N. (2010). Endemic diseases and african economic growth: challenges and policy responses. *Journal of African Economies 19*(suppl 3), 81–109.

Werker, E., A. Ahuja, and B. Wendell (2006). *Male circumcision and AIDS: the macroeconomic impact of a health crisis*. Division of Research, Harvard Business School.

Worobey, M., M. Gemmel, D. E. Teuwen, T. Haselkorn, K. Kunstman, M. Bunce, J.-J. Muyembe, J.-M. Kabongo, R. M. Kalengayi, V. M. Eric, G. M. T. P., and W. S. M. (2008). Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. *Nature 455*, 661–664.

Young, A. (2005). The gift of the dying: The tragedy of AIDS and the welfare of future African generations. *The Quarterly Journal of Economics 120*(2), 423–466.

Appendices

Table 8: Descriptives Region

	Regions	Mean	Sd	Min	Max
HIV_{total}	241	.055	.065	.001	.289
HIV_{male}	241	.042	.053	0	.231
HIV_{female}	241	.065	.075	.001	.337
$MaleCircumcision$	241	.578	.334	0	1
$TravelTimeKinshasa$	241	1.687	.575	.005	3.497
$Muslimshare$	236	.329	.371	0	1
$Ethnicity_0$	241	.117	.239	0	1
$Ethnicity_1$	241	.213	.302	0	1
$Ethnicity_2$	241	.188	.290	0	1
$Ethnicity_3$	241	.195	.352	0	1
$Ethnicity_4$	241	.0005	.006	0	.092

Table 9: Robustness - No Country Dummies (First Stage)
Dependent Variable: HIV Prevalence

	Model 1	Model 2
Initial Income	0.290	-0.164
	[0.489]	[0.450]
Male Circumcision	-0.076***	-0.035**
	[0.010]	[0.015]
Travel Time (in days)	0.003	0.030***
	[0.006]	[0.008]
Muslim Share		-0.083***
		[0.017]
Capital Region Indicator		0.025*
		[0.015]
Ethnicity Controls	No	Yes
Time Fixed Effects	Yes	Yes
Country Fixed Effects	No	No
Regions	241	236
F-Statistic	12.712	15.040
R-squared	0.306	0.459

Notes: This table shows the first stage regression estimates of male circumcision and distance to the first outbreak on HIV prevalence rates across 25 countries in sub-Saharan Africa for the period 2003 – 2012 without country fixed effects. Controlling for initial income, Muslim shares and ethnicity levels, overall the table establishes a statistically significant negative effect of male circumcision on HIV while the effect of distance becomes positive. Standard errors in brackets, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Robustness - No Country Dummies
Dependent Variable: Output Growth Per Capital

	OLS	2SLS	OLS	2SLS
HIV Prevalence	-0.310**	-0.539	-0.480***	-0.926*
	[0.152]	[0.470]	[0.167]	[0.559]
Initial Income	-2.319***	-2.293***	-2.424***	-2.654***
	[0.678]	[0.725]	[0.739]	[0.882]
Muslim Shares			-0.057*	-0.087*
			[0.030]	[0.048]
Capital Region Indicator			-0.034*	-0.024
			[0.019]	[0.023]
Ethnicity Controls	No	No	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
Country Fixed Effects	No	No	No	No
Regions	241	241	236	236
F-Statistic	10.896	11.859	11.374	10.045
R-squared	0.300	0.293	0.361	0.342
Cragg-Donald or Kleibergen-Paap		26.184		20.118
p-value Anderson-Rubin F-test		0.219		0.148
p-value Hansen J statistic		0.086		0.243

Notes: This table shows the second stage regression estimates of HIV prevalence – instrumented by male circumcision and distance to the first epidemic outbreak – on economic growth across 25 countries in sub-Saharan Africa for the period 2003 – 2012. Economic growth indicator is output per capita growth proxied by average growth in night light per capita. Controlling for initial income, Muslim shares and ethnicity levels but not controlling for country fixed effects, overall, the table establishes a negative (potentially significant) effect of the HIV prevalence rates on the growth of output per capita when ignoring country time-invariant effects. Standard errors in brackets, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Robustness - Alternative Origin (first stage)
Dependent Variable: HIV Prevalence

	Model 1	Model 2
Initial Income	0.213*	0.058
	[0.108]	[0.114]
Male Circumcision	-0.021***	-0.020**
	[0.008]	[0.009]
Travel Time Oster (in days)	-0.020**	-0.009
	[0.008]	[0.006]
Muslim Share		-0.027***
		[0.010]
Capital Region Indicator		0.014**
		[0.006]
Ethnicity Controls	No	Yes
Time Fixed Effects	Yes	Yes
Country Fixed Effects	Yes	Yes
Regions	241	236
F-Statistic	68.659	69.836
R-squared	0.869	0.883

Notes: This table shows the first stage regression estimates of male circumcision and distance to the first outbreak on HIV prevalence rates across 25 countries in sub-Saharan Africa for the period 2003 – 2012 when using the origin of Oster (2012). Controlling for initial income, Muslim shares and ethnicity levels, overall the table establishes a statistically significant negative effect of male circumcision on HIV while the effect of distance less significant. Standard errors in brackets, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 12: Robustness: Alternative Origin
Dependent Variable: Output Per Capita Growth

	OLS	2SLS	OLS	2SLS
HIV Prevalence	-0.347	-1.247	-0.490*	-2.020
	[0.259]	[0.970]	[0.286]	[1.729]
Initial Income	-2.971***	-2.685***	-2.746***	-2.560***
	[0.911]	[0.868]	[0.957]	[0.910]
Muslim Share			-0.034	-0.087
			[0.047]	[0.079]
Capital Region Indicator			-0.031	-0.011
			[0.019]	[0.029]
Ethnicity Controls	No	No	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
Country Fixed Effects	Yes	Yes	Yes	Yes
Regions	241	241	236	236
F-Statistic	11.118	8.918	8.425	7.608
R-squared	0.435	0.416	0.465	0.419
Cragg-Donald or Kleibergen-Paap		6.008		4.066
p-value Anderson-Rubin F-test		0.331		0.329
p-value Hansen J statistic		0.407		0.317

Notes: This table shows the second stage regression estimates of HIV prevalence – instrumented by male circumcision and distance to the first epidemic outbreak – on economic growth across 25 countries in sub-Saharan Africa for the period 2003 – 2012 when using the origin of Oster (2012). Economic growth indicator is output per capita growth proxied by average growth in night light per capita. Controlling for initial income, Muslim shares and ethnicity levels, overall, the table establishes an insignificant negative effect of the HIV prevalence rates on the growth of output per capita. Standard errors in brackets, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 13: Robustness: Gender Specific HIV (First Stage)
Dependent Variable: Female & Male HIV Prevalence

	HIV_{female}	HIV_{female}	HIV_{male}	HIV_{male}
Initial Income	0.416***	0.257**	0.027	-0.133
	[0.134]	[0.128]	[0.100]	[0.114]
Male Circumcision	-0.026***	-0.023**	-0.018**	-0.016*
	[0.010]	[0.010]	[0.007]	[0.009]
Travel Time (in days)	-0.024***	-0.014*	-0.013**	-0.005
	[0.009]	[0.007]	[0.006]	[0.005]
Muslim Share		-0.030**		-0.023***
		[0.012]		[0.008]
Capital Region Indicator		0.015**		0.012**
		[0.007]		[0.006]
Ethnicity Controls	No	Yes	No	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
Country Fixed Effects	Yes	Yes	Yes	Yes
Regions	241	236	241	236
F-Statistic	109.500	138.042	34.781	32.158
R-squared	0.867	0.880	0.843	0.858

Notes: This table shows the first stage regression estimates of male circumcision and distance to the first outbreak on gender specific HIV prevalence rates across 25 countries in sub-Saharan Africa for the period 2003 – 2012. Controlling for initial income, Muslim shares and ethnicity levels, overall the table establishes a statistically significant negative effect of male circumcision on HIV and a negative but less significant effect of distance. In general results are stronger for female prevalence rates. Standard errors in brackets, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 14: Robustness - Gender Specific HIV
Dependent Variable: Output Per Capita Growth

	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS
HIV Prevalence - female	-0.236	-0.987	-0.347	-1.696				
	[0.229]	[0.772]	[0.249]	[1.379]				
HIV Prevalence - male					-0.471*	-1.482	-0.624**	-2.329
					[0.283]	[1.306]	[0.314]	[2.341]
Initial Income	-2.954***	-2.548***	-2.690***	-2.241**	-3.039***	-2.947***	-2.865***	-3.026***
	[0.909]	[0.915]	[0.957]	[0.992]	[0.920]	[0.876]	[0.968]	[0.976]
Muslim Share			-0.031	-0.085			-0.035	-0.083
			[0.047]	[0.076]			[0.047]	[0.084]
Capital Region Indicator			-0.032*	-0.012			-0.031	-0.012
			[0.019]	[0.026]			[0.019]	[0.032]
Ethnicity Controls	No	No	Yes	Yes	No	No	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regions	241	241	236	236	241	241	236	236
F-Statistic	11.398	9.032	8.514	7.988	10.971	8.999	8.375	6.915
R-squared	0.434	0.416	0.464	0.413	0.436	0.417	0.467	0.420
Cragg-Donald or Kleibergen-Paap		6.575		4.582		5.362		2.793
p-value Anderson-Rubin F-test		0.405		0.385		0.405		0.385
p-value Hansen J statistic		0.498		0.473		0.427		0.323

Notes: This table shows the second stage regression estimates of gender specific HIV prevalence – instrumented by male circumcision and distance to the first epidemic outbreak – on economic growth across 25 countries in sub-Saharan Africa for the period 2003 – 2012. Economic growth indicator is output per capita growth proxied by average growth in night light per capita. Columns 1 – 4 show estimates using female prevalence while Columns 5 – 8 use male. Controlling for initial income, Muslim shares and ethnicity levels, overall, the table establishes an insignificant negative effect of the HIV prevalence rates on the growth of output per capita regardless of gender. Standard errors in brackets, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 15: Robustness - Instuments Individually (First Stage)
Dependent Variable: HIV Prevalence

Instrument:	Male Circumcision		Distance	
Initial Income	0.346***	0.098	0.187*	0.081
	[0.096]	[0.112]	[0.108]	[0.123]
Male Circumcision	-0.024***	-0.021**		
	[0.008]	[0.009]		
Travel Time (in days)			-0.020***	-0.011*
			[0.007]	[0.006]
Muslim Share		-0.030***		-0.030***
		[0.010]		[0.010]
Capital Region Indicator		0.015***		0.011*
		[0.006]		[0.006]
Ethnicity Controls	No	Yes	No	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
Country Fixed Effects	Yes	Yes	Yes	Yes
Regions	241	236	241	236
F-Statistic	68.964	71.897	68.124	66.585
R-squared	0.864	0.882	0.864	0.880

Notes: This table shows the first stage regression estimates of first male circumcision and second distance to the first outbreak on HIV prevalence rates across 25 countries in sub-Saharan Africa for the period 2003 – 2012. Controlling for initial income, Muslim shares and ethnicity levels, overall the table establishes a statistically significant negative effect of male circumcision and distance on HIV also when estimated individually. Standard errors in brackets, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 16: Robustness - Individual Instruments
Dependent Variable: Output Per Capita Growth

	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS
HIV Prevalence	-0.347	-0.528	-0.490*	-1.243	-0.347	-1.832	-0.490*	-3.874
	[0.259]	[1.276]	[0.286]	[1.882]	[0.259]	[1.300]	[0.286]	[3.009]
Initial Income	-2.971***	-2.914***	-2.746***	-2.655***	-2.971***	-2.500**	-2.746***	-2.335**
	[0.911]	[0.857]	[0.957]	[0.894]	[0.911]	[0.997]	[0.957]	[1.074]
Muslim Share			-0.034	-0.060			-0.034	-0.151
			[0.047]	[0.082]			[0.047]	[0.118]
Capital Region Indicator			-0.031	-0.021			-0.031	0.013
			[0.019]	[0.030]			[0.019]	[0.044]
Ethnicity Controls	No	No	Yes	Yes	No	No	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regions	241	241	236	236	241	241	236	236
F-Statistic	11.118	10.546	8.425	7.929	11.118	8.044	8.425	7.725
R-squared	0.435	0.434	0.465	0.454	0.435	0.383	0.465	0.236
Cragg-Donald or Kleibergen-Paap		8.333		5.797		8.296		3.887
p-value Anderson-Rubin F-test		0.697		0.539		0.193		0.198

Notes: This table shows the second stage regression estimates of HIV prevalence – instrumented by first male circumcision (columns 2 and 4) and second distance (columns 6 and 8) to the first epidemic outbreak – on economic growth across 25 countries in sub-Saharan Africa for the period 2003 – 2012. Economic growth indicator is output per capita growth proxied by average growth in night light per capita. Controlling for initial income, Muslim shares and ethnicity levels, overall, the table establishes an insignificant negative effect of the HIV prevalence rates on the growth of output per capita regardless of the instrument used. Standard errors in brackets, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 17: Robustness - Extended Sample (First Stage)
Dependent Variable: HIV Prevalence

	Model 1	Model 2
Initial Income	0.183***	0.124***
	[0.042]	[0.044]
Male Circumcision	-0.019**	-0.021**
	[0.009]	[0.009]
Travel Time (in days)	-0.016**	-0.006
	[0.008]	[0.006]
Muslim Share		-0.026***
		[0.010]
Capital Region Indicator		0.014**
		[0.006]
Ethnicity Controls	No	Yes
Time Fixed Effects	Yes	Yes
Country Fixed Effects	Yes	Yes
Regions	367	362
F-Statistic	76.172	83.984
R-squared	0.869	0.882

Notes: This table shows the first stage regression estimates of male circumcision and distance to the first outbreak on HIV prevalence rates across 25 countries in sub-Saharan Africa for the period 2003 – 2012. The sample includes two surveys for countries with available data. Controlling for initial income, Muslim shares and ethnicity levels, overall the table establishes a statistically significant negative effect of male circumcision on HIV and a negative effect of distance. Clustered (regional level) standard errors in brackets, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 18: Robustness - Extended Sample
Dependent Variable: Output per Capita Growth

	OLS	2SLS	OLS	2SLS
HIV Prevalence	-0.311	-0.545	-0.306	-0.328
	[0.201]	[0.845]	[0.216]	[1.251]
Initial Income	-1.542***	-1.498***	-1.485***	-1.482***
	[0.409]	[0.397]	[0.339]	[0.341]
Muslim Share			-0.004	-0.004
			[0.028]	[0.050]
Capital Region Indicator			-0.044***	-0.044**
			[0.015]	[0.021]
Ethnicity Controls	No	No	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
Country Fixed Effects	Yes	Yes	Yes	Yes
Regions	367	367	362	362
F-Statistic	18.055	12.004	9.604	9.661
R-squared	0.430	0.429	0.457	0.457
Cragg-Donald or Kleibergen-Paap		3.893		3.384
p-value Anderson-Rubin F-test		0.310		0.639
p-value Hansen J statistic		0.164		0.329

Notes: This table shows the second stage regression estimates of HIV prevalence – instrumented by male circumcision and second distance to the first epidemic outbreak – on economic growth across 25 countries in sub-Saharan Africa for the period 2003 – 2012. The sample includes two surveys for countries with available data. Economic growth indicator is output per capita growth proxied by average growth in night light per capita. Controlling for initial income, Muslim shares and ethnicity levels, overall, the table establishes an insignificant negative effect of the HIV prevalence rates on the growth of output per capita. Clustered (regional level) standard errors in brackets, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 19: Robustness - Concurrency as Third IV (First Stage)
Dependent Variable: HIV Prevalence

	Model 1	Model 2
Initial Income	0.119	0.004
	[0.108]	[0.120]
Male Circumcision	-0.019**	-0.016*
	[0.008]	[0.009]
Travel Time (in days)	-0.020***	-0.012**
	[0.007]	[0.006]
Sex Partners	0.025**	0.019**
	[0.011]	[0.008]
Muslim Share		-0.022**
		[0.009]
Capital Region Indicator		0.011*
		[0.006]
Constant	0.096***	0.053***
	[0.019]	[0.012]
Ethnicity controls	No	Yes
Time Fixed Effects	Yes	Yes
Country Fixed effects	Yes	Yes
Regions	241	236
F-Statistic	69.027	66.451
R-squared	0.877	0.887

Notes: This table shows the first stage regression estimates of male circumcision, distance to the first outbreak and concurrency (proxied by the average number of sexual partners excluding the spouse) on HIV prevalence rates across 25 countries in sub-Saharan Africa for the period 2003 – 2012. Controlling for initial income, Muslim shares and ethnicity levels, overall the table establishes a statistically significant negative effect of male circumcision and distance on HIV and a statistically positive effect of concurrency. Standard errors in brackets, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 20: Robustness - Concurrency as third IV
Dependent Variable: Output per Capita Growth

	OLS	2SLS	OLS	2SLS	LIML	GMM
HIV Prevalence	-0.347	-1.329	-0.490*	-2.248	-2.310	-2.279
	[0.259]	[1.323]	[0.286]	[1.738]	[1.784]	[1.543]
Initial Income	-2.971***	-2.659***	-2.746***	-2.533***	-2.525***	-2.685***
	[0.911]	[0.999]	[0.957]	[0.957]	[0.960]	[0.915]
Muslim Share			-0.034	-0.095	-0.097	-0.100
			[0.047]	[0.074]	[0.075]	[0.071]
Capital Region Indicator			-0.031	-0.008	-0.008	-0.007
			[0.019]	[0.029]	[0.029]	[0.027]
Ethnicity Controls	No	No	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Regions	241	241	236	236	236	236
F-Statistic	11.118	8.780	8.425	7.603	7.607	7.640
R-squared	0.435	0.412	0.465	0.404	0.399	0.400
Cragg-Donald or Kleibergen-Paap		5.805		5.493	5.493	5.493
p-value Anderson-Rubin LR					0.764	
p-value Anderson-Rubin F-test		0.534		0.383	0.383	0.383
p-value Hansen J statistic		0.757		0.682	0.685	0.682

Notes: This table shows the second stage regression estimates of HIV prevalence - instrumented by male circumcision, distance to the first epidemic outbreak and concurrency (proxied by the average number of sexual partners excluding the spouse) - on economic growth across 25 countries in sub-Saharan Africa for the period 2003 – 2012. Economic growth indicator is output per capita growth proxied by average growth in night light per capita. Controlling for initial income, Muslim shares and ethnicity levels, overall, the table establishes an insignificant negative effect of the HIV prevalence rates on the growth of output per capita. Standard errors in brackets, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

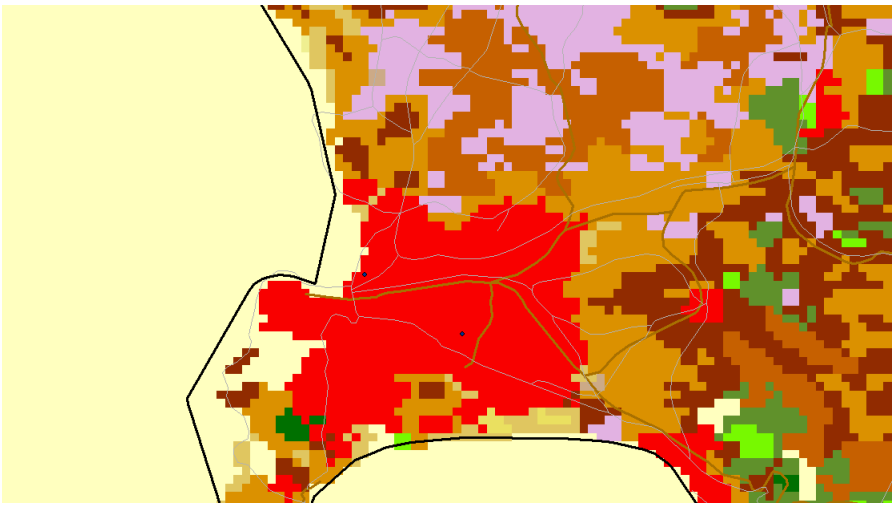
Estimating Travel Times

To estimate average travel times to Kinshasa, we follow Nelson (2008) and calculate the "transportation surface". We construct it by overlaying all the potential transportation possibilities and allocate speed to each mode of transportation. We also allocate speed to the base map by adjusting for terrain slope and altitude to account for slower or on foot transportation in hilly terrain and at higher altitudes. Figure 5 illustrates the method.

The colors in the background are the land cover data: the red is urban area, the green colors are various types of forests, the brown colors are woodlands (the darker the denser), the purple areas are crop lands, and the sandy color is grasslands. To each of these we allocate speed (i.e., how fast you can walk in these areas). Our speed allocation follows Nelson (2008)[48]. If there are no roads or other means of transportation in an area, walking speed will be the highest possible speed. If a road is on top of a woodland cell the speed will not be limited to the woodland walking speed (gray: roads, brown:

[48] http://bioval.jrc.ec.europa.eu/products/gam/index.htm

Figure 5: Transportation Surface



Source: Authors' calculations based on data from Nelson (2008)

rails). For each cell we end up with the maximum speed possible in that area (the transportation surface)[49]. We add a country-pair specific friction to crossing borders. In general crossing borders is time consuming and this friction reflects that. Finally, we calculate the minimum travel time to Kinshasa for each cell and take the average for each region[50]. Since regions' population density is not uniform, we also calculate the population weighted (based on LandScan 2000 data) average travel time as a robustness check of our baseline distance measures.

[49]The land cover data is from 2000 whereas the data on roads and rails are all from the mid-1990s. Even though the data are after the spread of the epidemic, we conjecture that current travel times should be correlated with historical travel times (though there may be some endogeneity); the robustness check with the great circle "as the crow flies" distances helps address this potential issue

[50]We do not speculate in the number of hours travelled per day and hence we implicitly assume travelling for 24 hours each day.