Understanding the Southern African ‘Anomaly’: Poverty, Endemic Disease and HIV

Larry Sawers and Eileen Stillwaggon

ABSTRACT

The epicentre of the global HIV epidemic is southern Africa. Previous explanations point to migration patterns and highly skewed income distribution, both thought to promote risky sexual behaviour. This study emphasizes the importance of common infectious and parasitic diseases that increase the likelihood of HIV transmission by increasing contagiousness and vulnerability to infection. Using multiple regression analysis on country-level data, the authors find that socio-economic variables explain statistically only one-tenth of the difference in HIV prevalence between southern Africa and other low- and middle-income countries. Measures of five cofactor infections together with the socio-economic variables, however, explain statistically about two-thirds of the southern Africa difference in HIV prevalence. They conclude that the relative affluence of countries in southern Africa and historical migration patterns have tended to mask the vulnerability of the majority of their populations who are poor and who have very high prevalence of infectious and parasitic diseases. Those diseases replicate a cycle of poverty that produces biological vulnerability through coinfections. An important implication of this research is that integrating treatment of endemic diseases with other HIV-prevention policies may be necessary to slow the spread of HIV.

INTRODUCTION

Average adult HIV prevalence in 2007 was 18 per cent in the nine countries of southern Africa (South Africa, Swaziland, Lesotho, Zimbabwe, Botswana, Mozambique, Namibia, Zambia and Malawi). Elsewhere in sub-Saharan Africa, HIV prevalence averaged 2.7 per cent in 2007; in other low- and middle-income countries, HIV prevalence averaged 0.5 per cent. The epicentre of the global HIV epidemic is southern Africa. Understanding the nature of HIV requires that we understand the region where the epidemic is most intense.
HIV is an opportunist virus that may have found in southern Africa a laboratory with the most propitious combination of factors promoting transmission. Various explanations have been proposed for the region’s vulnerability to HIV, but the present study is the first to attempt a comprehensive account of the southern Africa anomaly and to test it statistically. Explaining HIV in southern Africa may give us a key to understanding how social, biological and ecological factors interact to produce catastrophic results. That is particularly instructive for understanding a pathogen that has had relatively little impact in most other parts of the world. We examine the relevance of numerous factors considered in the literature and then construct statistical models of HIV prevalence. We find that, controlling for a combination of socio-economic and biological factors, the levels of HIV infection in southern Africa are not anomalous. They conform to what is known about HIV and reveal important vulnerabilities generally overlooked in discussions of southern Africa.

BACKGROUND: HISTORY, ECONOMICS AND HEALTH IN SOUTHERN AFRICA

Some of southern Africa’s historical features are unique, but many of its socio-economic, ecological and health characteristics are similar to those in the rest of Africa or in other low- and middle-income countries. The severity of the HIV epidemic in southern Africa is most probably the result of a combination of attributes it shares with other poor countries, especially in Africa, and features unique to southern Africa. The literature on HIV and southern Africa suggests a number of factors that might have aggravated the spread of HIV in the region.

Labour Migration and Apartheid

Perhaps the most obvious characteristic of southern Africa that sets it apart from the rest of the world is its history of apartheid and the earlier system of internal and international labour migration that this reinforced. Even before the National Party decreed strict residential segregation in South Africa in 1948, cross-border and internal labour migration supplied the labour for the country’s mines and factories. An extensive literature dating back to the 1940s addressed the health effects of migration in southern Africa, including the spread of tuberculosis and sexually transmitted infections (STIs) from migrants to their home communities (see Horwitz, 2001 for a summary of the early literature).

There is also a substantial literature that points to the role of migration in the spread of HIV, and many of those works focus on southern Africa in
High prevalence among migrants seems to support the assumption that long separation of spouses in circular migrant streams, single-sex hostels and higher levels of commercial sex work around mines and factories promote the spread of HIV (Lurie, 2004; see also Brummer, 2002; Clark et al., 2007; and Kanyenze, 2004, although without data). Migration has also been associated with HIV transmission in other parts of the world. In Ecuador, for example, over 80 per cent of the people infected with HIV in the first decade of the epidemic had worked in the United States or were partners of migrants to the USA. Circular migration between the USA and the Dominican Republic has also been associated with the spread of HIV in the Dominican Republic (Stillwaggon, 2006). Not all studies, however, have found that labour migrants are more likely to be HIV-infected. In Southeast Asia, Oppenheimer and colleagues found that labour migrants were more risk-averse and less likely to be infected (Oppenheimer et al., 1998, cited in Skeldon, 2002: 3).

Labour migration has played a very important role in the economies of southern Africa and very likely has aggravated the HIV epidemic. But migration is not unique to the region and may be insufficient to explain the difference in prevalence between southern Africa and West Africa, the Caribbean, the Middle East or Southeast Asia. In combination with other characteristics of southern Africa, however, migration may have been an important factor in the complex aetiology of the epidemic and should be included in a model of HIV transmission.

In the post-apartheid years, there have been significant changes in the southern African economies. Migration within South Africa and to South Africa from most neighbouring countries has increased, but the composition of the migrant labour force has changed significantly. Women are migrating in increasing numbers, both with and without male partners. Rising levels of male unemployment have changed the support relationships in the region (Posel and Casale, 2003). Mark Hunter argues that the rapid increase and enormous scale of the HIV epidemic in South Africa (and one might extend that analysis to other countries in the region) was not an inevitable result of a history of apartheid, although he acknowledges that labour migration and the enforcement of family separation in apartheid did set the stage. Rather, the decision to institute market-oriented reforms and austerity measures in the 1990s was disastrous for a large proportion of the population. The unemployment rate has risen to over 40 per cent, marital rates have fallen sharply and inequality has increased substantially in the last two decades (Hunter, 2007). Women’s migration has increased and informal settlements have mushroomed around both large and small cities (Crush et al., 2005; 1).
Posel, 2003; Posel and Casale, 2003). Hunter draws much-needed attention to the gap in health care provision that mirrors the stark inequality in income and wealth in South Africa (Hunter, 2007). (For more on inequality in health care, see Booyseon, 2002, 2003).

Cross-border and internal migration is socially disruptive and affects health in numerous ways. The literature on HIV and migration in southern Africa, however, focuses exclusively on the effect of migration on risky sexual behaviour, often without data, to the exclusion of other health-related factors (Brummer, 2002; Kanyenze, 2004). While the literature of social medicine in South Africa of the last half century has always included some mention of the miserable working and living conditions and inadequate nutrition of migrants (Horwitz, 2001; Marks, 2002), the sound bite that has emerged in AIDS policy discourse emphasizes ‘the constant movement of large numbers of sexually active men from town to countryside and back’ (Marks, 2002: 18). Certainly the behavioural implications of separating men from their families and communities were very important in the spread of STIs, and later HIV, in southern Africa. Nevertheless, that emphasis on sexual behaviour reinforces the prevailing explanation for high HIV prevalence in sub-Saharan Africa that focuses almost exclusively on risky sexual behaviour.

The literature on HIV and migration in southern Africa overlooks the direct biological impact of deprivation on migrant health and its role in accelerating the HIV epidemics in the region. There is a substantial body of evidence that poor sanitation, nutrition, and inadequate health care...

2. Poor sanitation spreads numerous bacterial, viral and parasitic infections. Borkow and Bentwich (2006) survey a large number of studies and trials investigating the connection between helminth infections and cell-mediated immunity. They argue that the majority of observational studies provide evidence for interaction between HIV and helminth infection, but there are many studies that do not show a connection and a few that suggest a protective effect of certain helminth infections. Almost all of those studies, however, are observational studies. Walson, Herrin and John-Stewart (2009) analysed almost 7,000 studies of HIV patients treated with anti-helminthics. They find only three randomized controlled trials, all three of which confirm the HIV–helminth connection. See also Bentwich et al. (2008); Chan (2005); Fawzi and Hunter (1998); Friis and Michaelson (1998); John et al. (1997); Landers (1996); Montresor et al. (2001); Nacher (2002); PPC (2002); Semba et al. (1994); Walson et al. (2008); Wolday et al. (2002); Woodward (1998). Additional studies are cited in Stillwaggon (2006). In contrast to other helminths, the hypothesized link between HIV and *S. hematobium* has to do with its effect on epithelial integrity and not only on its effect on cell-mediated immunity. See the in-text discussion of *S. hematobium* below. In our statistical analysis presented below, we do include any measure of helminth infection other than schistosomiasis.

3. Protein, fats, calories and most of the 40-plus micronutrients necessary for human health affect or modulate the immune system in some way and thus affect a person’s vulnerability to infectious disease. There is a ‘vast’ scientific literature showing that malnutrition undermines the immune system and that vitamin supplementation improves immune response (Mehta and Fawzi, 2007: 358). HIV is an infection and thus poor nutrition is likely to make people more susceptible to HIV infection. Malnutrition may also shorten the lives of...
for people living in single-sex barracks and squatter settlements\(^4\) can increase the transmission of HIV. A fully specified model of the HIV epidemics in southern Africa must include variables that reflect what is known about the interactions between HIV and other infectious and parasitic diseases.

**Poverty in Poor Countries**

From the beginning of the HIV epidemics in sub-Saharan Africa, researchers noted that HIV prevalence in many countries was higher in upper income and wealth brackets. (Gillespie, Kadiyala and Greener, 2007 is the most comprehensive survey of that literature; other reviews of the literature include Gillespie and Suneetha, 2005; Greener, 2004.) Most research on the connection between income, wealth and HIV looks at a single country or district within a country, but Mishra et al. (2007) use nationally representative surveys from eight sub-Saharan African countries to examine the issue. They find that HIV prevalence is higher in the wealthiest quintile than in the poorest quintile in all eight countries, usually by a substantial margin. Since antiretroviral treatment in poor countries is more readily available for the affluent than for poor people, wealthy people infected with HIV live longer on average and that biases comparisons of prevalence across wealth levels. Mishra et al. (2007) were not able to control for this effect, so the higher HIV prevalence of the wealthy is at least in part a result of the greater longevity of the wealthy rather than a reliable indicator of cumulative incidence. Moreover, wealthier people are more likely to live in cities and be better educated, both of which are associated with higher HIV prevalence. Wealthier men in the eight countries were also more likely to report two or more sexual partners in the previous year (except in Tanzania), more lifetime sexual

\[\text{those already infected with HIV and make them more contagious (Chan, 2005; Gillespie and Suneetha, 2005; Stillwaggon, 2006). The most comprehensive and up-to-date survey of studies and trials that examine the nutrition–HIV interaction (Mehta and Fawzi, 2007) sums up the evidence as follows: vitamin A supplementation reduces morbidity and delays mortality for HIV-infected children; multivitamin supplementation reduces adverse pregnancy outcomes in HIV-infected women, reduces mother-to-child transmission for poorly nourished mothers, and slows the progression to AIDS for HIV patients. On the other hand, the evidence from trials is that vitamin A supplementation alone does not protect against mother-to-child transmission of HIV except for pre-term babies (see Coutsoudis et al., 1999: 1522; Fawzi et al., 2001 provides corroborating evidence) and may even increase transmission (possibly due to dosing with massive amounts of the wrong form of the vitamin). The evidence about whether vitamin A protects adults from contracting HIV is mixed. In the statistical analysis below, we do not include any measure of nutrition because of the lack of suitable measures at the national level.}\]

4. Poor housing cannot provide adequate protection from disease vectors such as mosquitoes that spread malaria and flies that spread trachoma. The connections between HIV and malaria, and HIV and trachoma are discussed below.
partners, and sex with non-regular partners. On the other hand, the wealthy in the surveys are more likely to report having information about HIV and using condoms. Using a multiple regression analysis and controlling for these and other factors, Mishra et al. (2007) find that the wealth difference in HIV prevalence becomes statistically insignificant for both men and women in all but one country. Furthermore, there is some evidence that HIV incidence in the region is falling among the wealthy and better educated (Gillespie, Kadiyala and Greener, 2007: S12; Lopman et al., 2007). Nevertheless, we do not make a case that HIV is primarily a disease of poverty since, even with the above adjustments to the data, it is clear that it is not mostly the poor who contract HIV in poor countries.

The argument that this article makes, however, is that HIV is a disease of poor countries because widespread poverty increases prevalence for both poor and wealthy people. It is widely believed — though there is little direct evidence — that poverty can force women into transactional sex and into sexual relations where they do not have the power to negotiate safe sexual practices (Gillespie, Kadiyala and Greener, 2007: S6). Moreover, affluent men can afford to pay for the sexual services of poor and desperate women. The evidence from Mishra et al. (2007) is that the wealthy are more likely to engage in sexual relations outside of monogamous unions. In countries with more desperately poor women, both well-to-do men and poor women may be more likely to become infected with HIV. Members of both groups can then pass HIV on to their regular sex partners. If that were the case, we would expect poor countries to have higher HIV prevalence than wealthier countries.

Poor countries are also less likely to provide safe and effective medical care than higher income countries. Higher income people (especially women), better-educated people and people living in cities have greater access to physicians, clinics and hospitals, possibly explaining the higher HIV prevalence among those groups. There is substantial evidence that effective infection-control procedures (such as sterilizing syringes or specula) are not followed as consistently in poor countries as in higher income countries. In the countries that Mishra et al. (2007) studied, other than

5. There is substantial evidence that an important proportion of the HIV infections in sub-Saharan Africa and elsewhere result from iatrogenic/nosocomial transmission. For example, see Deuchert and Brody (2006, 2007a); Chan (2005); Gisselquist (2008). In addition, most of Volume 20 of the International Journal of STD and AIDS (2009) is devoted to the topic. Potterat (2009) authors the review article that leads the volume. Southern Africa seems to be unusually derelict in providing safe medical care, even when compared with other sub-Saharan African countries. Five of the nine countries in southern Africa use no autodisable syringes in child vaccination programmes. Elsewhere in sub-Saharan Africa, only 15 per cent of countries fail to do so (see Table 5). Another important reason for an association between poor health care and HIV is that diseases that may promote HIV transmission, such as schistosomiasis, malaria and STIs, are less likely to be treated or treated effectively and safely in countries without good health care.
Lesotho, wealthy men were more likely to be circumcised than poor men. Circumcision is thought to reduce HIV transmission, but in a poor country where unsafe medical practices are widespread, circumcision can spread HIV rather than protect against it. In sum, iatrogenic/nosocomial transmission of HIV will be high among the poor and perhaps even higher among the affluent within poor countries when compared with wealthy countries.

Poverty may push some people into risky sexual behaviour or into contact with unsafe and inadequate medical practices, but it also forces people into other behaviours that pose risks for their health. Poor countries on average have a higher share of the population that is poorly nourished, lives in dwellings that do not protect against disease vectors, and resides in communities with inadequate sanitation. Poor people thus are burdened with high rates of infectious and parasitic diseases. The health of the affluent in poor countries is also affected since they are surrounded by sick people and thus more likely to become ill than in countries where disease is less prevalent. For example, malaria cases exceeded 25 per cent of the population in twenty-seven African countries in 2006; schistosomiasis prevalence was 25 per cent or more in twenty-two African countries (WHO, 2008b, 2009). Living in a country where one is surrounded with disease raises the likelihood that any mosquito bite or contact with ground water will bring an infection. To the extent that some of those diseases promote the transmission of HIV, HIV is a disease of poor countries.

Inequality

All southern African countries except Malawi have above average income inequality. Namibia has the highest income inequality in the world and Botswana is third highest. South Africa, Lesotho, Zambia, Swaziland and Zimbabwe have Gini coefficients above 50, well above the average in the rest of Africa and in other low- and middle-income countries. There is a widely noted correlation across countries between income inequality and HIV prevalence (Gillespie et al., 2007; Mahal, 2002; Over, 1998; Stillwaggon, 2000, 2002; Tsafack and Bassolé, 2006), which appears to be related to the foregoing arguments about the connection between poverty and HIV. For a country at any given level of income, a higher level of income inequality means that the poor are a greater share of the population. If a large proportion of poor people in a country is an important driver of HIV epidemics, then a country with high income inequality should have a higher HIV prevalence. South Africa, Botswana, Namibia and Swaziland are middle-income countries (and Lesotho is an enclave surrounded by South Africa) and those countries have the highest HIV prevalence in the world. At the most fundamental level, in the four middle-income countries of southern Africa, stark inequality means that the relatively high per capita GDP masks the profound
poverty of the majority of the population and the direct impact of that poverty on health.

**Gender Inequality**

Another dimension of inequality is gender inequality. Many researchers argue — though with little direct evidence — that the oppression of women, by reducing their incomes, assets and power, makes it difficult for them to negotiate safe sexual practices and encourages them to engage in transactional sex (Brummer, 2002; Gillespie, Kadiyala and Greener, 2007; Hunter, 2007; Over, 1998; Tsafack, 2006). In that case, a country with greater gender inequality would have higher rates of HIV. On the other hand, greater gender equality could mean fewer traditional restrictions for women on sexual behaviours, some of which might be risky. At any rate, the evidence shows that variations in sexual behaviour between countries do not explain the higher prevalence of HIV in sub-Saharan Africa.

**Risky Sexual Behaviours**

The previous sections present the arguments that long-distance labour migration, poverty, income inequality and gender inequality could all lead to higher prevalence of risky sexual behaviours. As more data from survey research have become available, it has been increasingly clear — by almost all measures of risky sexual behaviour, including having multiple sexual partners — that sub-Saharan Africa is little different from other regions around the world. The countries with the highest measured rates of risky sexual behaviour are in North America and Western Europe, where HIV prevalence is very low, not in southern Africa, where HIV prevalence is very high (see Stillwaggon, 2006; Wellings et al., 2006).

Confronted with that evidence, defenders of the notion that risky sexual behaviour accounts for the higher HIV prevalence in sub-Saharan Africa have narrowed their argument to a single kind of sexual behaviour — overlapping ongoing relationships or concurrency (see especially Epstein, 2007; Halperin and Epstein, 2004, 2007; Mah and Halperin, 2008, 2009). Nevertheless, the supporters of the concurrency hypothesis have not shown that concurrency is unusually common in sub-Saharan Africa generally or southern Africa specifically, that HIV spreads faster with concurrency compared with other high-risk forms of multiple sexual relationships, or that the prevalence of concurrency is in fact correlated with HIV prevalence (Kretzschmar et al., 2009; Lurie, 2009; Lurie, Rosenthal and Williams, 2009; Stillwaggon and Sawers, 2010).

The supporters of the concurrency hypothesis argue that concurrency is far more common in sub-Saharan Africa than in other regions, but they give data
for only a few African countries and even fewer countries elsewhere.\(^6\) Some of the surveys they present as evidence of concurrency do not even measure concurrency, they misreport data, they cite studies that actually show low rates of concurrency in Africa, and none of their information on attitudes toward concurrency is about non-African countries and is thus irrelevant to a hypothesis that is essentially comparative (Stillwaggon and Sawers, 2010).

Mishra and Bignami-Van Assche (2009) compile all DHS surveys that measure concurrency using the same definition. In sub-Saharan Africa (ten countries for men and eight countries for women), 8.7 per cent of men and 0.8 per cent of women had had concurrent partners in the previous year (ibid., derived from Tables 5, 8 and 9). In contrast, in the USA, 11 per cent of men had had a concurrent relationship in the previous year and 12 per cent of women had had a concurrent relationship in the previous five years (Adimora, Schoenbach and Doherty, 2007; Adimora et al., 2002). The most comprehensive study of concurrency shows that it is not especially high in sub-Saharan Africa.\(^7\)

Concurrency cannot explain the higher prevalence of HIV in sub-Saharan Africa unless it is more dangerous than other forms of multiple partnering found elsewhere. Supporters of the concurrency hypothesis point to one mathematical model of disease transmission, which shows that certain kinds of concurrent relationships under certain conditions can spread HIV infections far faster than serial monogamy (Kretzschmar and Morris, 1996; Morris and Kretzschmar, 1997). The model, however, makes several crucial but completely unrealistic assumptions: that women and men are equally likely to have concurrent partners, that all partners have sexual contact with every partner every day, that partnerships on average last only a few months, and that each sexual contact has a 5 per cent chance of transmitting HIV. Every survey of concurrency for women shows that women are far less likely to have concurrent partners than men. When Morris and Kretzschmar (2000) recalibrate their model using actual survey data on male and female concurrency in Rakai, Uganda (and use realistic assumptions about the length of relationships), they find that concurrency leads to only 26 per cent more cases of HIV when compared with serial monogamy (not exponentially larger HIV prevalence). They achieve that result even though their Rakai model still uses the preposterous assumption of daily sex with all partners.

---

6. Halperin and Epstein (2004, 2007) and Mah and Halperin (2008, 2009), relying mostly on data from one city in Latin America, one country each in North America and Europe, three countries and two cities in Asia, repeatedly state that concurrency is less common outside of Africa.

7. The supporters of the hypothesis dismiss the study because it is based on household surveys (although most of the data they offer in support of their hypothesis are also from household surveys) and because the surveys do not measure concurrency (although they do measure concurrency) (Mah and Halperin, 2009; Stillwaggon and Sawers, 2010).
The Rakai model also continues to assume a per-act transmission rate of .05 (from a study of Thai soldiers, nearly half of whom had STIs). Transmission among otherwise healthy adults is about one fiftieth of that rate, but transmission rates are far higher if either partner has STIs, schistosomiasis and/or malaria (Boily et al., 2009; Chan, 2005; Deuchert and Brody, 2007b). Table 5 shows that the burdens of all of those diseases are far higher in sub-Saharan Africa than elsewhere. The model produces the rapid spread of HIV with concurrency by assuming transmission rates that are only found among unhealthy people. In other words, it is not concurrency that is spreading HIV in sub-Saharan Africa; it is cofactor infections that, by promoting the transmission of HIV, are driving the HIV epidemics in the region.

Even if concurrency was both an especially dangerous form of multiple partnering and also more prevalent in southern Africa than elsewhere, it would be difficult to argue that concurrency explains the high prevalence of HIV in the region without demonstrating that the prevalence of HIV and concurrency are indeed correlated. The risk posed by concurrency is not what one does, but what one’s partner is doing (which is often unknown), so the only direct way to test the concurrency hypothesis is to map sexual networks. The network study offered as evidence for the concurrency hypothesis (Helleringer et al., 2009) maps a sexual network, but it cannot show a statistically significant correlation because only 22 per cent of persons interviewed were willing to be tested for HIV and because it failed to distinguish between serially monogamous and concurrent relationships. Moreover, the study was conducted on a remote island and may tell us little about sexual networks elsewhere. A correlation within a population between self-reported concurrency (not that of their partner) and HIV prevalence cannot provide direct evidence, but at least that correlation would be consistent with the concurrency hypothesis. There have been four such ecological studies and none of them finds a statistically significant correlation (Jewkes et al., 2006; Lagarde et al., 2001; Mattson et al., 2007; Mishra and Bignami-Van Assche, 2009).

In short, the concurrency hypothesis is built on unconvincing empirical evidence and mathematical models that make critically important but completely unrealistic assumptions. Efforts to establish a statistical correlation between concurrency and HIV have failed. The stress on concurrency in sub-Saharan Africa is another stage in the two-decades-old practice of explaining HIV epidemics in Africa without empirical evidence, driven only by presumptions about the exceptional character of African sexuality.

---

8. Mah and Halperin (2009) also argue that two surveys (unnamed and without citation) of newly infected people thought their partners had other partners. Those findings tell us only that the respondents believed the message they repeatedly heard from HIV-prevention programmes for adults, which talk only about safe sex. Those who believe the message and have HIV will surely suspect that their partners are having sex with someone else.

9. See Stillwaggon (2003) for an explanation of why social scientists and policy makers have assumed, without evidence, that African sexuality is exceptional.
Cofactor Infections

Between otherwise healthy adults in developed countries, HIV has very low rates of heterosexual transmission (Chan, 2005; World Bank, 1997). In the developing world, including southern Africa, however, the majority of the population is not ‘otherwise healthy’. Field studies and trials as well as laboratory research have found substantial evidence for a connection between the transmission of HIV and several diseases highly endemic in southern Africa. The relative affluence of southern Africa and the emphasis on risky sexual behaviour have led most policy makers to overlook the biological vulnerability of the majority of southern Africa’s population to infectious disease.

There seems to be little doubt that sexually transmitted infections (STIs) increase transmission of HIV. Some STIs lead to genital ulceration creating entry points for the HIV virus, thereby making those who are not infected with HIV more vulnerable to infection and making infected persons more likely to transmit HIV. Both ulcerative and non-ulcerative STIs also make those who are not infected with HIV more vulnerable to HIV infection by producing genital inflammation, which draws CD4+ cells to the infected tissues. Those cells are the channel through which HIV infects the body. Furthermore, both ulcerative and non-ulcerative STIs make those who are infected with HIV more contagious by increasing viral shedding, which is the discharge of virus in the genital tract. In poor populations, bacterial STIs are more likely to be untreated because of lack of access to health care (Sturm et al., 1998). The health burden of untreated STIs is far higher in African countries than in other low- and middle-income countries (see Table 5). Together, these facts suggest that one of the reasons for the high prevalence of HIV in southern Africa is the high prevalence of untreated STIs.

Urinary schistosomiasis (S. hematobium), which afflicts more than 100 million people in sub-Saharan Africa (and almost nowhere else; WHO, 2008a), also can act as a cofactor of HIV transmission much in the same way as do STIs. Worms and ova of S. hematobium infect the reproductive tract

---

10. See Fleming and Wasserheit (1999) for their review of more than eighty trials; see also Boulton and Gray-Owen (2002); Chan (2005); Corbett et al. (2002); Grosskurth et al. (1995); Hayes et al. (1995); Laga et al. (1993); Rotchford et al. (2000); Wang et al. (2001); Wolday et al. (2004).

11. Trachoma might seem unrelated to HIV since it is an eye infection, but trachoma is caused by the same bacteria (C. trachomatis) as the STI chlamydia. As many as 70 per cent of women and 25 per cent of men with chlamydia are asymptomatic (Webster et al., 1993). That suggests the possibility that trachoma sufferers are coinfected genitally without the knowledge of the patient or medical personnel. Alternatively, high prevalence of trachoma may be merely a proxy for poor health care, since it is easily treated in its early stages and extracts a very high toll in blindness that would not be tolerated in a country with an adequate health system. Consequently, we include the trachoma variable in some of the regressions presented below because we think it suggests an important factor worthy of further investigation.
of both men and women, producing lesions and inflammation of the genital area (Attili et al., 1983; Feldmeier et al., 1995; Leutscher et al., 1998; Marble and Key, 1995). The genital ulcers of *S. hematobium* are indistinguishable without biopsy from those of ulcerative STIs. A recent trial in Zimbabwe found that genital lesions of schistosomiasis increased HIV risk in women three-fold compared to women in the same villages without genital lesions of schistosomiasis (Kjetland et al., 2006). The WHO Department of Control of Neglected Tropical Diseases recently identified the interactions of HIV with endemic parasitic and infectious diseases as a priority for research (WHO, 2007: 45), but AIDS policy organizations have not adopted protocols for treating coinfections of these locally endemic diseases.

Malaria, which is especially prevalent in sub-Saharan Africa, interacts with HIV, increasing viral load up to ten-fold. Viral load remains elevated in HIV-infected persons for six weeks after malarial episodes (Abu-Raddad et al., 2006; Hoffman et al., 1999; Kublin et al., 2005). Elevated viral load not only increases individual contagiousness of HIV-infected persons, it also affects the dynamics of the epidemic at a population level (Abu-Raddad et al., 2006; Bloland et al., 1995; Corbett et al., 2002; Herrero et al., 2007; Hoffman et al., 1999; Whitworth et al., 2000; Xiao et al., 1998; for additional discussion and sources, see Stillwaggon, 2006). Individuals in malaria-endemic areas have a higher probability of sexual contact with persons who have high viral load due to coinfection with malaria and who are therefore more contagious. In the AIDS literature, there is increasing and warranted attention to the first few months after infection with HIV because people have high viral loads at that time and are thus more likely to transmit HIV to their partners (Brenner et al., 2007; Cohen and Pilcher, 2005; Pilcher et al., 2004; Quinn et al., 2000). What also needs to be considered is that malaria and other coinfections produce recurring elevation of viral loads over the lifetime of the HIV-infected person, not just in the first few months following HIV infection.

12. A recent study shows that rhesus monkeys infected with *S. mansoni* are more susceptible to SIV (simian immunodeficiency virus) infection and have elevated viral load compared to monkeys without schistosomiasis (Chenine et al., 2008). Human trials with *S. mansoni* have not looked for differential incidence (whether coinfection with schistosomiasis makes one more vulnerable to HIV) but have examined viral load and CD4+ counts of coinfected HIV patients (that is, whether HIV patients are more contagious) and how they are affected by treatment with antihelminthics. Those trials have produced mixed results. Kallestrup et al. (2005) support the hypothesis that schistosomiasis coinfection makes HIV patients more contagious by raising viral load, but others do not (Brown et al., 2005; Elliott et al., 2003; Lawn et al., 2000; Mwinzi et al., 2001). There are 54 million people in sub-Saharan Africa afflicted with *S. mansoni*. The argument of this article is that the genital manifestations of *S. hematobium* make it different from other species of schistosomiasis as well as other helminths.

13. We have provided a rationale for a long list of independent variables in our model, but we should point out a variable that we are not including. Numerous colleagues have suggested we add clade to our analysis. The subtype of HIV-1 that is dominant in southern Africa
METHOD

This research extends the work of authors who use cross-national data and multiple regression analysis to estimate the effect of factors thought to drive the HIV epidemic (Bonnel, 2000; Deuchert and Brody, 2007a; Drain et al., 2004; Mahal, 2002; Over, 1998; Stillwaggon, 2000; Stillwaggon, 2002; Tsafack, 2008). All but one of those studies focus on a specific issue that the author(s) believe especially important: risky sexual behaviours, gender discrimination, human development, income inequality, economic growth, iatrogenic transmission of HIV, and nutrition. Those authors all use the same or similar socio-economic control variables and then add other variables relevant to the specific issue addressed in their studies. The most common control variables used in those studies are income inequality, per capita income (level or growth rate), urbanization (level or growth rate), proportion of the population that is Muslim, age of the epidemic, a measure of gender discrimination, literacy or school enrolment rates, and one or more binary variables for region. Deuchert and Brody (2007a) examined the iatrogenic or nosocomial spread of HIV. Their study found a measure of unsafe injection practices to be positively and significantly correlated with HIV prevalence, so we also include that variable in our equations.

To those control variables we add one not used in previous studies. We include a measure of cross-border migration because of the special importance of migration in the discussion of HIV in southern Africa. We call all of these control variables ‘the basic model’. Finally, in the present study, which extends our previous research (Sawers, Stillwaggon and Hertz, 2008), is clade C, but that subtype is relatively rare in most other countries (the most important exceptions being Bangladesh, Papua New Guinea, India and Nepal). That fact suggests to many observers that clade C is more fit for transmission than other clades and several contributors to the virological literature have speculated just that (Chan, 2005; Laurent et al., 2007; Ping et al., 1999; Rainwater et al., 2005; Spira et al., 2003; Tebit et al., 2007). There is, however, little evidence to confirm that suspicion. One study finds that vaginal shedding in pregnant women is greater with clade C (John-Stewart et al., 2005), but another study found no difference in mother-to-child transmissibility of the virus among different clades (Tapia et al., 2003). A third study finds no evidence of different fitness for heterosexual transmission of the different clades (Morison et al., 2001). The most complete data on HIV clades are from the HIV Sequence Database provided by the Los Alamos National Laboratory at http://www.hiv.lanl.gov/content/sequence/HIV/mainpage.html, from which one can obtain the per cent of HIV infections of clade C for seventy-five of the ninety-one countries in our dataset. That variable is strongly and positively correlated with both HIV prevalence and the southern Africa variable in our dataset, but when the variables of the basic model are added to the regression, the regression coefficient on clade C becomes unexpectedly negative and significant at only the 96 per cent level. If one then adds the cofactor infections variables to the regression, the regression coefficient on the clade C variable loses its statistical significance (the t statistic is only 0.40, significant at the 0.691 level). Accordingly, our research provides no support for the notion that differences in HIV sequences between countries can explain the higher HIV prevalence in southern Africa using multivariate analysis.
we add to the basic model measures of cofactor infections that are thought to promote the transmission of HIV.

The regional binary variable in our model takes the value of one for the nine countries in southern Africa and zero for other countries. A positive and significant coefficient on the southern Africa variable tells us that HIV prevalence is higher in southern Africa than elsewhere, but does not tell us why. Our objective is to explain differences in HIV prevalence among low- and middle-income countries (not just find statistical associations and a high $R^2$). Location is not a policy-sensitive variable. Although using a regional binary variable raises the predictive value of the model in a technical sense, it begs the question of what is causing high prevalence in the region. Accordingly, we measure our success in explaining the variation in HIV among countries by how far we can reduce the size of the regression coefficient on the southern Africa variable.

Using data for ninety-one developing and transition countries, we estimate a number of ordinary least squares multivariate regressions on the log of adult HIV prevalence in 2007. All of the regressions reported here include the variables of the basic model: a binary variable for the nine countries of southern Africa, the Gini coefficient (the most widely used measure of income inequality), log of per capita income on a purchasing power parity basis, per cent of adults who are literate, per cent of the population living in urban areas, age of the epidemic, a measure of gender discrimination (one minus the ratio of female economic participation to male economic participation), per cent of the population who are Muslim, workers’ remittances as a percent of GDP, and a binary variable that takes the value of one if the country used no autodisable syringes in child vaccination programmes. Table 1 lists the source for each of the variables. The first step is to regress

14. Countries with per capita income in 2003 of less than US$ 12,000 on a purchasing power parity basis were included in the analysis. Rather than imputing missing values, we dropped countries with missing data from the analysis, leaving a dataset of ninety-one countries. See Appendix for a list of countries in the dataset. The present study follows established practice by using the log form of HIV prevalence, per capita income and, where possible, measures of cofactor infections, which reduces the influence of outliers and generally leads to more robust and efficient estimates when analysing highly skewed variables such as these.

15. It would have been preferable to use HIV incidence to describe how the epidemic is unfolding. Prevalence includes historical information and is confounded by greater survival in countries with effective antiretroviral treatment programmes, but incidence data are generally not available.

16. All measures of literacy and school enrolment for our sample are highly correlated and perform similarly as regressors. In our sample, the adult literacy rate and the female-to-male literacy ratio are virtually identical in a statistical sense (with a simple correlation of 0.99). Using the latter variable, however, gives the misleading impression that we are using it to measure gender discrimination; hence the present study uses the more straightforward variable, adult literacy, to avoid that confusion.
Table 1. Sources of Data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult HIV prevalence, 2006</td>
<td>Per cent of persons aged 15 to 49 with HIV</td>
<td>UNAIDS (2008)</td>
</tr>
<tr>
<td>Gini coefficient</td>
<td>Measure of income inequality, 2004</td>
<td>World Bank, World Development Indicators</td>
</tr>
<tr>
<td>Income per capita</td>
<td>Per capita GDP (PPP basis), 2003</td>
<td>World Bank, World Development Indicators</td>
</tr>
<tr>
<td>Urbanization</td>
<td>Per cent of population in urban areas, 2003</td>
<td>World Bank, World Development Indicators</td>
</tr>
<tr>
<td>Age of the epidemic</td>
<td>Years since first diagnosis of HIV</td>
<td>World Bank (1997: 318–24)</td>
</tr>
<tr>
<td>Adult literacy</td>
<td>Per cent of adult population who are literate, 2003</td>
<td>UNDP, Human Development Indicators, 2005</td>
</tr>
<tr>
<td>Female participation ratio</td>
<td>One minus ratio of female economic activity rate to male economic activity rate, 2005</td>
<td>UNDP, Human Development Indicators, 2007</td>
</tr>
<tr>
<td>Muslim per cent</td>
<td>Per cent of the population who are Muslim</td>
<td>CIA World Factbook</td>
</tr>
<tr>
<td>Remittances</td>
<td>Workers’ remittances as a per cent of GDP</td>
<td>World Bank, World Development Indicators</td>
</tr>
<tr>
<td>Autodisable syringe use</td>
<td>Did not use autodisable syringes in child vaccination programmes</td>
<td>WHO (2005: Annex Table 7)</td>
</tr>
<tr>
<td>STIs, schistosomiasis, trachoma</td>
<td>DALYs (Disability-adjusted life years) for each disease per 100,000 people, 2002</td>
<td>WHO (2004)</td>
</tr>
<tr>
<td>Malaria prevalence</td>
<td>Cases as a per cent of population</td>
<td>WHO (2008b)</td>
</tr>
</tbody>
</table>

only the variables of the basic model on HIV prevalence. Next, we add each cofactor infection variable, one at a time, to a regression with the basic model variables. Lastly, we regress all of the cofactor infections together with the basic model variables. Estimation of all of the regressions is robust to heteroskedasticity.

RESULTS

In the equation presented in Table 2, only the variables of the basic model are regressed on HIV prevalence. The $R^2$, that is, the per cent of the variance in the log of adult HIV prevalence associated with the independent variables taken together, is 64.0 per cent. Six of the ten independent variables are significant at the 95 per cent level or better and another is significant at the 90 per cent level. The southern Africa variable, the Gini coefficient, and the age of the epidemic are positively correlated with HIV prevalence. The Muslim per cent, adult literacy, the gender discrimination index and income per capita variables are negatively associated with HIV prevalence. The regression coefficient on each independent variable in the model tells us the change in the log of HIV prevalence — after accounting for the effects
Table 2. Multiple Regression Analysis of HIV Prevalence with Basic Model and without Cofactor Infections (N = 91)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>t Statistic</th>
<th>Probability</th>
<th>Beta Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Southern Africa</td>
<td>2.7901</td>
<td>6.35</td>
<td>0.000</td>
<td>.4717</td>
</tr>
<tr>
<td>Gini coefficient</td>
<td>.0291</td>
<td>1.71</td>
<td>0.091</td>
<td>.1521</td>
</tr>
<tr>
<td>Income per capita (log)</td>
<td>−.4862</td>
<td>−2.01</td>
<td>0.047</td>
<td>−.2478</td>
</tr>
<tr>
<td>Urbanization</td>
<td>.0083</td>
<td>0.87</td>
<td>0.387</td>
<td>.0951</td>
</tr>
<tr>
<td>Age of epidemic</td>
<td>.1003</td>
<td>2.41</td>
<td>0.018</td>
<td>.1777</td>
</tr>
<tr>
<td>Adult literacy</td>
<td>−.0189</td>
<td>−2.16</td>
<td>0.034</td>
<td>−.2551</td>
</tr>
<tr>
<td>Female participation ratio</td>
<td>−.0195</td>
<td>−2.51</td>
<td>0.014</td>
<td>−.2047</td>
</tr>
<tr>
<td>Muslim</td>
<td>−.0089</td>
<td>−2.56</td>
<td>0.012</td>
<td>−.1926</td>
</tr>
<tr>
<td>Remittances</td>
<td>.0083</td>
<td>0.50</td>
<td>0.618</td>
<td>.0294</td>
</tr>
<tr>
<td>Autodisable syringe use</td>
<td>.2527</td>
<td>0.94</td>
<td>0.351</td>
<td>.0715</td>
</tr>
<tr>
<td>Constant</td>
<td>−.2009</td>
<td>−0.10</td>
<td>0.919</td>
<td></td>
</tr>
</tbody>
</table>

of the other independent variables — when the value of the independent variable changes by 1. Since the southern Africa variable takes the value of either 1 or 0, its coefficient tells us the difference in the log of HIV prevalence between southern Africa and elsewhere after taking into account the effects of the other basic model variables.

Average HIV prevalence in southern Africa is about 16 times the average prevalence in other low- and middle-income countries in our dataset. The regression coefficient on the southern Africa variable is 2.79 in the regression using only the variables in the basic model. To find the ratio of HIV prevalence in southern Africa and elsewhere that is not explained by the basic model variables, one takes the antilog (with an adjustment for heteroskedasticity) of the coefficient on the southern Africa variable. The regression coefficient on southern Africa tells us that HIV prevalence in southern Africa not explained by the model is 14.9 times the level in other low- and middle-income countries. In other words, the basic model ‘explains’ in a statistical sense less than 10 per cent of the nearly 16-fold differential in HIV prevalence between the southern African and other low- and middle-income countries.

17. The coefficient on the southern Africa variable estimates the difference between the mean log of HIV prevalence in southern Africa compared to other countries in the dataset. In order to convert this to a statement about the difference in mean HIV prevalence (as opposed to mean log prevalence), one must take the antilog of that coefficient, but also correct for the non-zero mean of the antilogged error terms. That correction takes into account the heteroskedasticity in error terms between southern Africa and the other countries and yields a close approximation of the actual ratio of means. The ratio of the predicted means is the antilog of [the regression coefficient on the southern Africa variable plus (half the difference between the variance of the regression residuals in southern Africa and the variance of the regression residuals outside of southern Africa)]. Thomas Hertz devised this calculation.
Table 3. Multiple Regression Analysis of HIV Prevalence with Basic Model and each Cofactor Infection Separately (N = 91)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>t Statistic</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chlamydia (log)</td>
<td>.8932</td>
<td>4.33</td>
<td>0.000</td>
</tr>
<tr>
<td>Gonorrhoea (log)</td>
<td>.8894</td>
<td>5.70</td>
<td>0.000</td>
</tr>
<tr>
<td>Syphilis (log)</td>
<td>.4426</td>
<td>5.37</td>
<td>0.000</td>
</tr>
<tr>
<td>STIs (log)</td>
<td>1.0076</td>
<td>6.21</td>
<td>0.000</td>
</tr>
<tr>
<td>Schistosomiasis</td>
<td>.0074</td>
<td>4.21</td>
<td>0.000</td>
</tr>
<tr>
<td>Malaria</td>
<td>.0689</td>
<td>5.89</td>
<td>0.000</td>
</tr>
</tbody>
</table>

The next step in the analysis is to add each of the five cofactor infection variables — one at a time — to a multiple regression in which the other independent variables are those in the basic model. Schistosomiasis, chlamydia, syphilis and gonorrhoea are measured by DALYs per 100,000 population, and malaria is measured by the annual number of cases as a per cent of the population. The results of those five regressions are reported in Table 3. All of the cofactor infection variables are positively correlated with HIV prevalence, and all of the cofactor infection variables are significant at the 99.9 per cent level.

The regressions reported in Table 3 do not measure the true importance of each cofactor infection since they are subject to missing variable bias, that is, they do not include the other cofactor infections that are correlated with HIV prevalence. On the other hand, the five cofactor infections are highly collinear. (The simple correlation \([R]\) in the sample between each pair among the five cofactor infections ranges between 0.56 and 0.90 — and exceeds 0.68 for all pairs excluding malaria.) We entered a variety of different combinations of the five diseases into regressions with the basic model variables, but they produced different results depending on which cofactor infections were entered and how they were combined. Collinearity prevents us from determining the separate effect of each infection. Nevertheless, the objective of our research is not to show the effect of any particular disease, but to show that the burden of certain diseases taken together explains statistically the variations in HIV prevalence across countries. Accordingly, we add together the four cofactor infections that are measured the same way (in DALYs per 100,000 population) and are thus additive. We call this variable ‘combined cofactors variable’. (Malaria is entered separately.

18. Because the vast majority of deaths from malaria are in infants, the DALYs associated with malaria are extremely high. That measure, however, does not reflect the burden of malaria among sexually active adults, unlike the DALYs from other diseases. We therefore use malaria prevalence as our measure.

19. For the same reason (statistical efficiency) that we use the log of HIV prevalence and per capita income, we also use the log of chlamydia, syphilis and gonorrhoea in these regressions. There are many countries in the dataset with no schistosomiasis or malaria and one cannot calculate a logarithm of zero.
since it is measured differently.) Combining the four cofactor infections into a single variable eliminates the collinearity problem, but also makes epidemiological sense since it measures the combined burden of the four cofactor infections, all of which are thought to promote HIV transmission through the same mechanisms. All southern African countries have very high levels of chlamydia, gonorrhoea and schistosomiasis (110 or more DALYs per 100,000 population) and only African countries have very high levels of those three infections. Five southern African countries also have very high levels of syphilis and only three countries outside Africa have levels that high.

Accordingly, we estimate a regression with the basic model variables together with the malaria prevalence and the combined cofactors variable. That raises the $R^2$ from 66.8 to 80.1 per cent and raises the $F$ statistic for the equation from 22.3 to 55.1. Both the combined cofactors variable and malaria prevalence are significant at the 99.9 per cent level and the $F$ statistic for those two variables taken together is 42.1, significant at the 99.99 per cent level. That test confirms our hypothesis that cofactor infections are significantly correlated with HIV prevalence after controlling for other variables. The beta coefficients for the two variables measuring cofactor infections are larger than for any other variable (.490 and .442) and substantially larger than either the southern Africa variable (.320) or the per cent Muslim variables (−.206). The beta coefficients measure which independent variables have a greater statistical impact on the dependent variable after taking into account the different units by which the variables are measured.

The regression coefficient on the southern Africa variable in the equation including the cofactor infection variables is 1.89, which tells us that the HIV prevalence not ‘explained’ by the other variables in the equation is 5.5 times that of other low- and middle-income countries. In other words, adding the cofactor infections to the model allows us to ‘explain’ in a statistical sense about 70 per cent of the 16-fold differential in HIV prevalence between southern Africa and other low- and middle-income countries, substantially more than the model without the cofactor diseases, which explains less than 10 per cent of the southern Africa differential.

One can compare the regression coefficient and $t$ statistic on the log of per capita income in the regression with only the variables of the basic model (in Table 2) to those statistics in the regression that includes the cofactor

20. As noted above, there is a substantial medical literature that supports our contention that STIs and *S. hematobium* are cofactors with HIV, but there is no such literature connecting trachoma to HIV. Nevertheless, it is plausible that persons infected with trachoma also have unrecognized genital infection with the same bacteria, that is, they have chlamydia. We re-estimate the regression presented in Table 4, replacing the combined cofactors variable with one that includes trachoma. The two regressions are almost identical except that when trachoma is included, the coefficient on the southern Africa variable falls from 1.89 to 1.76. In other words, including trachoma enhances our ability to ‘explain’ the southern Africa differential.
### Table 4. Multiple Regression Analysis of HIV Prevalence with Basic Model and Cofactor Infections (N = 91)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>t Statistic</th>
<th>Probability</th>
<th>Beta Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combined cofactors</td>
<td>.7127</td>
<td>4.61</td>
<td>0.000</td>
<td>.4905</td>
</tr>
<tr>
<td>Malaria prevalence</td>
<td>.0492</td>
<td>3.74</td>
<td>0.000</td>
<td>.4423</td>
</tr>
<tr>
<td>Southern Africa</td>
<td>1.8924</td>
<td>5.94</td>
<td>0.000</td>
<td>.3200</td>
</tr>
<tr>
<td>Gini coefficient</td>
<td>.0284</td>
<td>2.09</td>
<td>0.040</td>
<td>.1487</td>
</tr>
<tr>
<td>Income per capita (log)</td>
<td>.2932</td>
<td>1.58</td>
<td>0.118</td>
<td>.1494</td>
</tr>
<tr>
<td>Urbanization</td>
<td>−.0023</td>
<td>−0.27</td>
<td>0.788</td>
<td>−.0259</td>
</tr>
<tr>
<td>Age of epidemic</td>
<td>.0274</td>
<td>0.90</td>
<td>0.373</td>
<td>.0485</td>
</tr>
<tr>
<td>Adult literacy</td>
<td>.0134</td>
<td>1.81</td>
<td>0.074</td>
<td>.1808</td>
</tr>
<tr>
<td>Female participation ratio</td>
<td>−.0075</td>
<td>−1.01</td>
<td>0.315</td>
<td>−.0787</td>
</tr>
<tr>
<td>Muslim</td>
<td>−.0095</td>
<td>−3.41</td>
<td>0.001</td>
<td>−.2055</td>
</tr>
<tr>
<td>Remittances</td>
<td>.0436</td>
<td>2.73</td>
<td>0.008</td>
<td>.1536</td>
</tr>
<tr>
<td>Autodisable syringe use</td>
<td>.3165</td>
<td>1.38</td>
<td>0.173</td>
<td>.0896</td>
</tr>
<tr>
<td>Constant</td>
<td>−10.4735</td>
<td>−5.90</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>

Infections in Table 4. The coefficient on per capita income in the regression without the cofactor infection variables is −0.486 (significant at the 95 per cent level), that is, lower income countries, other things being equal, have higher HIV prevalence. Adding the cofactor infection variables to the equation turns the coefficient on per capita income positive and eliminates its statistical significance. That is consistent with our argument that what is important about the inverse association between HIV prevalence and per capita income is the high disease burden of the poor.

### Limitations

The statistical analysis presented above shows that after controlling for a number of factors including the presence of cofactor infections, location in southern Africa by itself accounts for less than a third of the 16-fold difference in HIV prevalence between that region and other low- and middle-income countries in our sample. Two points should be emphasized. That the coefficient on the southern Africa variable remains significantly positive means that our model suffers from missing variable bias. The southern Africa variable is without causal content. If it remains statistically significant in the regression equations, it means that something else sets southern Africa apart from the rest of the world that we have not yet discovered or found a way to measure adequately. Some examples of variables that one would like to add to the equation include condom use, a better measure of injection safety practices in medical and non-medical settings, various measures of risky sexual behaviour,\(^{21}\) prevalence of circumcision (not just

---

\(^{21}\) The most widely available measure of risky sexual behaviour is Age at First Sex for Females, but that variable has not been measured for almost half of the countries in our
DISCUSSION

In retrospect, it is clear that, in southern Africa, HIV was an epidemic ‘waiting to happen’ (Marks, 2002). The factors that made southern Africa an ideal setting for HIV to flourish are numerous. What follows is a discussion of how southern Africa compares with the rest of sub-Saharan Africa and with other low- and middle-income countries. Table 5 presents means of the variables used in this analysis for four groups of countries — southern Africa, central Africa (the belt of a dozen contiguous countries that stretches across Africa from Angola to Kenya), the rest of sub-Saharan Africa, and other low- and middle-income countries outside sub-Saharan Africa.

Southern African shares many characteristics with the rest of sub-Saharan Africa that set them apart from other low- and middle-income countries. Excluding South Africa, the proportion of the population who live on less than US$ 2 a day (not shown in Table 5 because it was not used in the regressions) is the same in southern Africa (71 per cent) as elsewhere in sub-Saharan Africa. That compares with an average of 40 per cent in other dataset. Twenty-two of the forty-six countries for which there are such data are in Africa, so those forty-six countries are a representative sample of neither African countries nor low- and middle-income countries. Sexual behaviour surveys reflect a bias in data collection common in AIDS research. The assumption that the AIDS epidemics in sub-Saharan Africa result from something distinctive about African sexuality means that there are more surveys about sexual behaviour in African countries than in other developing countries. For a discussion of bias in data selection in Africa, see Stillwaggon (2003).

22. We know that HIV can spread from country to country across land borders. Truck routes are frequently mentioned as conduits for the spread of HIV in sub-Saharan Africa and prevalence is higher along highways than in more remote parts of the same country. In an attempt to capture those neighbourhood effects, we measure, for every country in the sample, the average HIV prevalence of all countries with which it shares a land border. Adding that variable to the regression in Table 4 reduces the coefficient on southern Africa to insignificance (the t statistic is 0.20 and the p is 0.789) while leaving the coefficients on the cofactor infection variables significant at the 98 per cent level. In that sense, we have completely explained the southern Africa differential in HIV prevalence. Nevertheless, we have merely shifted our ignorance back one step: in a statistical sense, we know why each country in southern Africa has a high level of HIV, but we have not explained why all of its neighbours have high prevalence. Furthermore, the southern Africa variable and the HIV prevalence in neighbouring countries variable are so highly correlated (R = 0.92) that they are practically the same variable. That collinearity makes it difficult to tease out the separate effects of the two locational variables. In short, there are some very complicated spatial dynamics to the spread of HIV across sub-Saharan Africa and around the world that cannot be usefully analysed with an OLS multiple regression on cross-section data.

23. The proportion of the population living on US$ 2 a day would be an appropriate variable to use in the regression analysis except for missing data. That measure was not available for thirty-two out of the ninety-one countries in the data set.
Table 5. Means of Variables in Low- and Middle-income Countries by Region

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sub-Saharan Africa</th>
<th>Other sub-Saharan Africa</th>
<th>Other Low- and Middle-income Countries</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Southern</td>
<td>Central</td>
<td></td>
</tr>
<tr>
<td>Chlamydia DALYs per 100,000</td>
<td>120 119 119</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>Gonorrhoea DALYs per 100,000</td>
<td>156 162 168</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>Syphilis DALYs per 100,000</td>
<td>120 285 521</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>Schistosomiasis DALYs per 100,000</td>
<td>238 236 277</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Trachoma DALYs per 100,000</td>
<td>149 149 187</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Combined cofactors (STIs and schistosomiasis)</td>
<td>638 809 1095 145</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malaria prevalence, % of population</td>
<td>12.8 32.4 32.2</td>
<td>0.9</td>
<td></td>
</tr>
<tr>
<td>HIV prevalence, % of adult population</td>
<td>18.3 4.3 1.72</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>Gini coefficient</td>
<td>54.1 42.3 40.2</td>
<td>41.5</td>
<td></td>
</tr>
<tr>
<td>Income per capita PPP in US$</td>
<td>4391 1729 1348</td>
<td>5139</td>
<td></td>
</tr>
<tr>
<td>Urbanization, per cent of population</td>
<td>33.8 39.0 33.0</td>
<td>51.9</td>
<td></td>
</tr>
<tr>
<td>Age of epidemic in years</td>
<td>21.6 21.2 20.2</td>
<td>19.6</td>
<td></td>
</tr>
<tr>
<td>Adult literacy, per cent of population</td>
<td>73.2 69.2 36.8</td>
<td>84.6</td>
<td></td>
</tr>
<tr>
<td>Female economic participation ratio</td>
<td>.721 .833 .715</td>
<td>.649</td>
<td></td>
</tr>
<tr>
<td>Muslim, per cent of population</td>
<td>9.8 16.0 58.9</td>
<td>24.8</td>
<td></td>
</tr>
<tr>
<td>Remittances as per cent of GDP</td>
<td>5.9 2.5 3.1</td>
<td>5.9</td>
<td></td>
</tr>
<tr>
<td>Autodisable syringe, per cent did not use</td>
<td>.556 .222 0.0</td>
<td>41.7</td>
<td></td>
</tr>
</tbody>
</table>


low- and middle-income countries. Consequently, the cofactor infections analysed in this study (either diseases of poverty or diseases less likely to be treated because of poverty) are found at roughly similar levels throughout most of sub-Saharan Africa, but at far lower levels outside Africa. STIs are epidemic around the world, but their burden in sub-Saharan Africa is more than triple the level in other low- and middle-income countries. Moreover, the difference for the other cofactor diseases is far greater. The burden of schistosomiasis, for example, is between 18 and 21 times as great in sub-Saharan Africa as it is in other low- and middle-income countries. Recall that *S. hematobium*, which affects the genital tract by creating lesions, inflammation and viral shedding in ways likely to promote HIV transmission, is found almost exclusively in Africa. To the extent that poverty and the diseases that poverty promotes lead to high HIV prevalence, then southern Africa is similar to the rest of sub-Saharan Africa and much more vulnerable than other low- and middle-income countries.

Other factors that could promote the spread of HIV set southern Africa apart from the rest of sub-Saharan Africa, but make it more similar to other developing and transition countries. A number of authors have noted the positive correlation of educational attainment and literacy with HIV (for example, Brent, 2006). Adult literacy (and the ratio of female to male literacy, which is not shown in Table 5) in southern Africa is a little lower than in low- and middle-income countries outside of sub-Saharan Africa.
and a little higher than in central Africa, but is much higher than in the rest of sub-Saharan Africa. In addition, the level of migration (as measured by remittances as a per cent of GDP) is the same in southern Africa as it is in low- and middle-income countries outside Africa and more than twice the level in the remainder of sub-Saharan Africa.

Income inequality puts southern Africa in a category by itself. Gini coefficients in most of southern Africa are among the highest in the world, but in the rest of sub-Saharan Africa, Gini coefficients average about the same as in low- and middle-income countries elsewhere. Furthermore, southern Africa on average has the lowest percentage of the population who are Muslim among the four groups of low- and middle-income countries in Table 5.

In sum, why has HIV reached its highest levels in southern Africa — one-quarter of the adult population in the most gravely affected countries? Perhaps what is unique about southern Africa is that it faces higher risk from so many of the factors that promote HIV epidemics. With regard to the level of poverty and the burden of cofactor infections, it is more like the rest of sub-Saharan Africa than countries elsewhere. With regard to literacy and migration, southern Africa is more like low- and middle-income countries outside of Africa than those within Africa. With regard to high levels of income inequality and low proportion of Muslims, southern Africa sets itself apart from the rest of the world. All of these factors come together in southern Africa to produce the most intense epidemics of HIV in the world.

CONCLUSIONS

AIDS prevention policy has failed to stem the spread of HIV in southern Africa and neighbouring countries. The research presented here suggests that policy makers have targeted too few factors in their attempts to slow or reverse the epidemics. The socio-economic variables in our basic model that are statistically associated with HIV prevalence, however, are not easily affected by changes in policy. Levels of inequality have been shown to be highly resistant in the short run to changes in the policy regime. Even the most successful efforts to promote economic development produce single-digit annual increases in per capita income. Migration patterns, literacy or female participation in the economy are slow to change. Moreover, after decades of trying to convince people to avoid risky sexual behaviour, there appears to be only scattered success.

In contrast, the variables correlated with HIV prevalence in this study that are the cheapest and easiest to change are the cofactor infections. Dramatic reductions in all of the cofactor infections can be produced rapidly and at low cost. Some of the infections can be cured by medicines that cost pennies per patient and can be distributed by schoolteachers or health care providers with little training. They are highly effective, can be easily tolerated and
are heat-stable with long shelf lives. Arguably, the easiest and fastest way to promote economic development, to alleviate poverty and even to slow or reverse the spread of HIV is to attack the cofactor infections analysed in this study. The present research provides support for integrating HIV prevention policy with efforts to address the multitude of health issues that beset Africa.

Treating or preventing diseases endemic in developing countries with antihelminthics, antibiotics and other interventions is a pragmatic, safe, ethically sound and relatively inexpensive strategy. Such interventions have substantial beneficial outcomes on their own, in better health, better school performance and higher productivity at work. Moreover, the improvement in overall well-being in itself promotes healthier behaviours in healthier people. Some HIV treatment protocols now include food for people taking antiretrovirals and treatment for tuberculosis because failing to do so wastes lives and resources. Nutritional supplementation for HIV-infected individuals may also reduce viral loads (making them less contagious), postpone the onset of AIDS, prevent opportunistic infections and prolong the efficacy of antiretrovirals. Similarly, our research supports the arguments of medical researchers that redirecting some of the funds designated for HIV prevention to deworming, sanitation, STI treatment, mosquito control and safe water supports, rather than undermines, HIV prevention, care and treatment. Treating locally endemic infections and the conditions that cause those infections in order to control HIV is not mission creep; rather it may be essential to the central mission of AIDS programming.

Appendix. Countries in Dataset

<table>
<thead>
<tr>
<th>Albania</th>
<th>Ecuador</th>
<th>Latvia</th>
<th>Philippines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algeria</td>
<td>Egypt</td>
<td>Lesotho</td>
<td>Poland</td>
</tr>
<tr>
<td>Armenia</td>
<td>El Salvador</td>
<td>Lithuania</td>
<td>Romania</td>
</tr>
<tr>
<td>Azerbaijan</td>
<td>Ethiopia</td>
<td>Madagascar</td>
<td>Russian Federation</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>Gabon</td>
<td>Malawi</td>
<td>Rwanda</td>
</tr>
<tr>
<td>Belarus</td>
<td>Gambia</td>
<td>Malaysia</td>
<td>Senegal</td>
</tr>
<tr>
<td>Benin</td>
<td>Georgia</td>
<td>Mali</td>
<td>Sierra Leone</td>
</tr>
<tr>
<td>Bolivia</td>
<td>Ghana</td>
<td>Mauritania</td>
<td>South Africa</td>
</tr>
<tr>
<td>Botswana</td>
<td>Guatemala</td>
<td>Mexico</td>
<td>Sri Lanka</td>
</tr>
<tr>
<td>Brazil</td>
<td>Guinea</td>
<td>Moldova</td>
<td>Swaziland</td>
</tr>
<tr>
<td>Bulgaria</td>
<td>Guinea-Bissau</td>
<td>Mongolia</td>
<td>Tajikistan</td>
</tr>
<tr>
<td>Burkina Faso</td>
<td>Guyana</td>
<td>Morocco</td>
<td>Tanzania</td>
</tr>
<tr>
<td>Burundi</td>
<td>Haiti</td>
<td>Mozambique</td>
<td>Thailand</td>
</tr>
<tr>
<td>Cambodia</td>
<td>Honduras</td>
<td>Namibia</td>
<td>Togo</td>
</tr>
<tr>
<td>Cameroon</td>
<td>India</td>
<td>Nepal</td>
<td>Trinidad and Tobago</td>
</tr>
<tr>
<td>Chile</td>
<td>Indonesia</td>
<td>Nicaragua</td>
<td>Tunisia</td>
</tr>
<tr>
<td>China</td>
<td>Iran</td>
<td>Niger</td>
<td>Turkey</td>
</tr>
<tr>
<td>Colombia</td>
<td>Jamaica</td>
<td>Nigeria</td>
<td>Uganda</td>
</tr>
<tr>
<td>Congo</td>
<td>Jordan</td>
<td>Pakistan</td>
<td>Ukraine</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>Kazakhstan</td>
<td>Panama</td>
<td>Viet Nam</td>
</tr>
<tr>
<td>Cote d’Ivoire</td>
<td>Kenya</td>
<td>Papua New Guinea</td>
<td>Yemen</td>
</tr>
<tr>
<td>Croatia</td>
<td>Kyrgyzstan</td>
<td>Paraguay</td>
<td>Zambia</td>
</tr>
<tr>
<td>Dominican Republic</td>
<td>Laos</td>
<td>Peru</td>
<td></td>
</tr>
</tbody>
</table>
REFERENCES


Poverty, Endemic Disease and HIV


Larry Sawers is Professor of Economics at American University in Washington, DC and is the corresponding author; he can be contacted at lsawers@american.edu. His research interests are in the field of economic development and he has published books and articles on Argentina, Ecuador, Tanzania and Lithuania. His research in recent years has focused on the HIV epidemics in Africa.

Eileen Stillwaggon is Professor of Economics and Harold G. Evans Eisenhower Professor at Gettysburg College in Pennsylvania. She has written a number of scholarly articles and books on poverty, health and development and on HIV epidemics in poor populations.