UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

WORKING PAPERS IN ECONOMICS

No 425

Economic Inequality and HIV in Malawi

Dick Durevall and Annika Lindskog

Revised November 2011

ISSN 1403-2473 (print)
ISSN 1403-2465 (online)

Department of Economics
School of Business, Economics and Law at University of Gothenburg
Vasagatan 1, PO Box 640, SE 405 30 Göteborg, Sweden
+46 31 786 0000, +46 31 786 1326 (fax)
www.handels.gu.se info@handels.gu.se

Economic Inequality and HIV in Malawi

Dick Durevall and Annika Lindskog

dick.durevall@economics.gu.se

annika.lindskog@economics.gu.se

Department of Economics,

School of Business, Economics and Law,

University of Gothenburg

P.O. Box 640, SE 405 30 Gothenburg, Sweden

Abstract

To analyze if the spread of HIV is related to economic inequality we estimate multilevel models of the individual probability of HIV infection among young Malawian women. We find a positive association between HIV infection and inequality at both the neighborhood and district levels, but no effect of individual poverty. We also find that the HIV-inequality relationship is related to risky sex, gender violence, and return migration, though no variable completely replaces economic inequality as a predictor of HIV infections. The HIV-inequality relationship does not seem to be related to bad health, gender gaps in education or women's market work.

JEL: I12.

Key words: Africa, AIDS, gender inequality, gender violence, Malawi, poverty.

*We would like to thank Daniela Andrén, Arne Bigsten, Heather Congdon, David Sahn, Måns Söderbom, Rick Wicks and anonymous referees for useful comments. We would also like to thank SIDA/SAREC for financial support.

1. Introduction

Poverty is typically viewed as an important driver of the HIV epidemic, and AIDS is often called a "disease of poverty".[1] However, several studies have recently shown that poor individuals are not more likely to be HIV positive than wealthy ones, and the poorest of the less developed countries do not have higher infection rates than other less developed countries (Gillespie et al., 2007; Whiteside, 2008, p. 53). Instead, economic inequality, together with gender inequality, has been suggested as a main socioeconomic driver of the spread of HIV (Nattrass, 2008; Whiteside, 2008, Ch. 3; Fox, 2010).

The idea that economic inequality and health are related is well-established. Since the beginning of the 1990s over 200 articles have been published on the topic, and though the results vary, many find a strong association between various health indicators and income equality across countries or regions within countries (Deaton, 2003, Subramanian and Kawachi, 2004; Wilkinson and Pickett, 2006). Yet, surprisingly few studies have analyzed economic inequality and HIV/AIDS and all seem to use cross-country data (Holmqvist, 2009; Tsafack Temah, 2009; Sawers and Stillwaggon, 2010a). Although useful, cross-country regressions are likely to suffer from omitted variable biases since many potentially relevant variables cannot be included. Moreover, if absolute income matters for health and there are diminishing health returns, a relationship between health and income inequality is produced at the aggregate level even though income inequality has no causal effect on health (Gravelle et al., 2002).

We analyze the association between economic inequality and HIV infections in Malawi; the country with the 9th highest national HIV rate in the world, at 11.0% in 2009 (UNAIDS, 2010). Since the size of the community might affect the results (Wilkinson and Pickett, 2006), two levels of community are included in the analysis: Malawi's 27 districts and the immediate neighborhood, measured by the sampling cluster used in the 2004 Malawi Demographic and Health Survey (MDHS). More specifically, we consider the effect of district consumption inequality and neighborhood wealth inequality on individual-level risks of HIV infection among Malawian women aged 15-24. The statistical analysis is carried out using multilevel logistic models of the probability of being HIV infected, combining data from the 2004 MDHS with district-level data from the 1997/98 Integrated Household and Income Survey and 1987 Population and Housing Census.

[1] See for example, Whiteside (2002), Fenton (2004), Stillwaggon (2006; 2009) and Wellings (2006).

We limit our sample to young women since they are likely to have been infected recently. This alleviates the potential problem of higher mortality among the poor, which affects studies including all prime-age adults (Sawers and Stillwaggon, 2010a). There are not enough HIV infected young men in the data to allow estimations on this group as the prevalence rate was only 2.1% (NSO and OCR Macro, 2005). The group of young women is also of particular interest since intergenerational transmission of HIV, which is sustaining the epidemic in the long run, mainly occurs via young women.

Our main findings are that there is a strong positive association between the risk of HIV infection and economic inequality. The relationship between HIV status and indicators of poverty, i.e., household wealth, neighborhood median wealth and district median consumption, is less clear-cut. There is no evidence that poorer women are more likely to be HIV positive than others, while the results for neighborhood and district-levels are mixed.

We also evaluate potential causes of the HIV-inequality relationship. The relationship appears to be due to risky sexual behavior and gender violence, which are more common in unequal societies, but not to indicators of bad health or gender gaps in education and women's market work. To some extent, the HIV-inequality relationship can be explained by high levels of return migration from urban to rural areas, which seem to affect both inequality and HIV prevalence rates. However, no variable completely replaces economic inequality as a predictor of HIV infections.

The paper is organized as follows. Section 2 briefly reviews earlier studies of the impact of poverty and inequality on HIV/AIDS. Section 3 describes the HIV epidemic in Malawi, and Section 4 presents our estimations strategy. Section 5 first describes the HIV data and possible sample election problems, and then the explanatory variables. Section 6 reports the empirical results, and Section 7 summarizes, discusses and concludes.

2. Inequality, Poverty and HIV/AIDS: What Do We Know?

In this section we first review the empirical evidence on HIV and economic inequality, poverty, and wealth. The focus is on Sub-Saharan Africa, where HIV mainly is transmitted through sexual contacts in the general adult population.[2] We then discuss mechanisms that potentially create links between economic inequality, poverty and HIV.[3]

There is strong empirical evidence that income inequality is associated with HIV prevalence at the country level. A recent contribution is Holmqvist (2009) who reviews other studies and carry out his own analysis. The Gini coefficient of income almost always has a statistically significant coefficient. Other studies that obtain similar results are Over (1998) Nattrass (2008), Tsafack Temah (2009) and Sawers and Stillwaggon (2010a). The size of the effect varies with specification, but a representative finding is that a change in the Gini coefficient from 0.4 to 0.6, roughly Malawi compared to South Africa, raises prevalence by 0.5 to 1 percentage point.

Studies analyzing poverty and HIV vastly outnumber those on inequality and HIV, and the findings are not as clear-cut. Cross-country analyses give mixed results when all countries (with available data) are included. When samples are restricted to developing countries, there is usually no impact of GDP per capita or poverty on the spread of HIV (Holmqvist, 2009). In fact, relatively rich African countries have higher infection rates than poor ones.

There are also various studies challenging the view that poor individuals have a higher risk of HIV infection (Lachaud, 2007; Mishra et al., 2007; Fortson, 2008; Msisha et al., 2008a). Using mainly DHS data for a number of Sub-Saharan countries, they often find that wealthy individuals are more or equally likely to be HIV positive. For example, Mishra et al. (2007) find that Malawian men in the three richest wealth quintiles are about 2.5 times more likely to be infected than those in the two poorest wealth quintiles. Wealthier individuals could have a higher infection risk due to riskier sexual

[2] The second most important channel is mother-to-child transmission of HIV. But we have data on HIV status in 2004 for women over 14 years, and people born with HIV 15 years earlier had already died by then. Some infections among adults are probably due to injections with unsterilized needles and blood transfusion with infected blood. Generally these channels are believed to be of minor importance compared to heterosexual contact, although there are divergent views (Stillwaggon, 2006; Mishra et al., 2008).

[3] There are innumerable studies of the causes of the HIV epidemic in general that are not covered here; Whiteside (2008) and UNAIDS (2010) provide general reviews.

behavior, but possibly also because of their greater access to health services that might expose them to the virus through injections, surgery, etc. (Mishra et al., 2008).

A possible caveat for these findings is that wealthier people might survive longer with HIV: in cross-sectional data HIV prevalence could then be higher for richer people even if the poor have higher or equal incidence rates (Gillespie et al., 2007). Lopman et al. (2007), using Zimbabwean panel data on incidence, show empirically that wealthy HIV-positive individuals have higher survival rates than poor HIV-positive individuals, particularly among men. However, summarizing the findings of Lopman et al. and two other recent panel data studies on HIV incidence (Bärnighausen et al., 2007; Hargreaves et al. 2007), there does not appear to be a systematic pattern between getting infected and individual income.

To the best of our knowledge, there are only two previous studies that analyze the role of poverty at the regional level within a country: Lachaud (2007) on Burkina Faso, and Msisha et al. (2008b) on Tanzania. They measure poverty by the headcount ratio and find it to be inversely related to HIV. Hence, several studies find that income inequality matters, while most studies on income and poverty, at individual, regional and country levels, fail to find support for the hypothesis that HIV is more common among the poor.

The association between income inequality and HIV prevalence raises questions about the mechanisms involved. In the literature on the relationship between income inequality and health in general, three main hypotheses have been suggested: the absolute income hypothesis, the relative income hypothesis, and the society-wide effects hypothesis (Leigh et al., 2009).

According to the absolute income hypothesis, it is really poverty, not income inequality, which generates the relationship. A region with high average income could have bad health when there is high income inequality simply because there are many with low incomes. Additionally, if there are diminishing health returns to income then an analysis of aggregate data produces a relationship between income inequality and health even though income inequality has no causal effect on health (Gravelle et al, 2002).[4] In other words, if income is transferred from a rich to a poor person, economic

[4] Though most empirical studies have considered the relationship between income and health, the absolute income effect should be due to consumption or permanent income; income matters since it determines consumption of health inputs such as nutritious food, health care and medication.

inequality is reduced, and health is improved as it declines less for the rich person than it improves for the poor person.

The relative income hypothesis states that income inequality is an indicator of social distance between individuals, and the larger the distance the more psychosocial stress and, consequently, the worse health (Wilkinson and Pickett, 2006).[5] If this is the case, health in a society can decline even if everybody receives a higher income. This would happen if there is an increase in inequality that increases the psychosocial stress among a large part of the population. Although the relative income hypothesis is most popular in social science fields outside of economics, the idea that 'utility' depends on comparisons of own income and consumption to that of others has been long established in economics (Veblen, 1899; Duesenberry, 1949). Recently this idea has gained considerable empirical support through studies in behavioral economics (Luttmer, 2005; Johansson-Stenman and Martinsson, 2006; Fliessbach et al., 2007).[6]

The society-wide effects are related to social capital, where inequality reduces trust and increases crime and violence (Leigh et al., 2009). This mechanism is related to the relative income hypothesis, since, for instance, low social status makes people feel disrespected, which in turn can generate violence (Wilkinson and Pickett, 2006). Another possible society-wide effect is lower provision of public goods since preferences and needs are likely to vary more in a more heterogeneous population (Banerjee and Somanathan, 2007). Farmer (1999) provides an example by documenting how tuberculosis, a preventable disease, can prevail among the poor in a rich and unequal society such as the US as the disease does not affect the large majority of the population.

There is little agreement on the relative importance of the three hypotheses. Wilkinson and Pickett (2006) and Babones (2008) conclude that there is ample support for the second and third hypotheses. Deaton (2003), on the other hand, argues that there is no direct link to ill health from income inequality; the empirical findings are due to factors other than income inequality per se, poverty being one explanation. Jen et al. (2009) obtain support for the poverty and diminishing health returns to income hypothesis, while Leigh et al. (2009) go even further, arguing that the relationship between

[5] For social distance, consumption, wealth or perhaps education inequalities might be at least as important as income inequalities. The economic literature on relative economic status, however, tends to emphasize income and consumption.

[6] A key part of the prospect theory of Kahnemann and Tversky (1979) is that the utility of an outcome depends on how it compares to some reference point.

income distribution and health is fragile or non-existent. However, they base their argument only on 'robustly estimated panel specifications' which might be too demanding if a change in inequality affects health with a long lag (Deaton, 2003; Glymour, 2008). Subramanian and Kawachi (2004) take the middle view, arguing that the results are inconclusive, although inequality seems to matter in unequal societies such as the U.S.

Since HIV is primarily transmitted through sexual intercourse in sub-Saharan Africa, the potential mechanisms that relate economic inequality to the spread of HIV might differ from those relevant for health in general. What matters for the spread of HIV is behavior, i.e. type and frequency of sexual contacts, and environment, i.e. susceptibility to the virus of uninfected people and the contagiousness of infected people. There is an on-going debate on whether risky sex, primarily concurrency, or unhealthy environment is the most important explanation for the high HIV rates in Southern and Eastern Africa. Concurrency implies that more people are connected in sexual networks at a given point in time, and there is a larger probability that a newly infected person has a sexual encounter with someone else soon after infection, when the viral load is high (Morris and Kretzscmar, 1997; Halperin and Epstein 2004; Mah and Halperin, 2010). Relatedly, transactional sex, i.e., the exchange of sex for material support, seems to be more common in Southern and Eastern Africa than elsewhere (UNAIDS, 2002; Hunter, 2002; Dunkle et al., 2004). Swidler and Watkins (2007) use data on everyday conversations in rural Malawi and find that transactional sex is common, more or less accepted, and related to concurrency.

However, the importance of concurrency is disputed. Sawers and Stillwaggon (2010b) and Lurie and Rosenthal (2010) argue that the empirical support is weak or non-existent, and Mapingure et al. (2010) fail to find that the number of sexual partners matters when comparing samples from Tanzania and Zimbabwe. Instead, undernourishment, infectious diseases, and bad health in general, are claimed to be more important intermediating factors, since they increase the per-contact transmission rate (Stillwaggon, 2006, 2009; Sawers and Stillwaggon, 2010a). Moreover, some sexually transmitted diseases and urinary schistosomiasis increase susceptibility through genital ulcers (Flemming and Wasserheit, 1999; Kjetland et al., 2006). Other diseases can also increase the viral load, and thus make HIV positive people more infectious; there is, for example, strong evidence that malaria increases the viral load in already infected people (Abu-Raddad, 2006).

The absolute income hypothesis is relevant for HIV/AIDS, since there is agreement that low income is related to poor health status in less developed countries, (Wilkinson and Pickett, 2006). As mentioned,

bad health could increase transmission rates. Moreover, according to economic theory, poverty could makes people short-sighted, and therefore more likely to take risks, since they care little about what happens to them ten years later (Oster, 2007). It has also been suggested that extreme poverty could induce women to exchange sex for goods or money to stay above the subsistence level (Swidler and Watkins, 2007; Tawfik and Watkins, 2007), while men could be induced to leave their families for extended periods to work far away from home, increasing the likelihood of extra marital affairs (Arrehag et al., 2006). Furthermore, poor people are more vulnerable to external shocks, such as drought, and the combined effect of poverty and shocks may increase risky behavior substantially (Bryceson and Fonseca, 2006).

It is also possible that a high level of poverty in a society increases infection risks for all, not only for the poor (Sawers and Stillwaggon, 2010a). If there is sexual networking between richer and poorer people, then undernourishment and an underfunded health care sector with unsafe practices could for example interact with transactional sex, putting both the poor and the non-poor at greater risk of being infected. This would not be captured by individual-level income, and could be the reason why studies fail to find that poverty matters; an analysis using the level of income in the community would however capture the effect.[7]

The main direct behavioral link between economic inequality and HIV is likely to be through transactional sex. In more unequal societies, relatively poor women may have sexual relationships because of aspirations to 'live a better life', not necessarily to secure the survival of themselves and their children (Fox, 2010). This seems to be the case in Malawi: Tawfik and Watkins (2007) find that women in rural areas engage in transactional sex, not mainly to secure subsistence living, but for attractive consumer goods. Moreover, in unequal societies there are likely to be more wealthy men that can afford transactional sex: in economics language, high inequality implies a low cost of an additional partner for wealthy men (Over, 1998).

Economic inequality could also increase the spread of HIV because of society-wide effects, notably due to lack of social cohesion (Barnett and Whiteside 2002, pp. 88-97). This could occur because it is difficult to mobilize collective action to implement effective responses to the epidemic in places with little social cohesion. There could also be more gender violence in more unequal societies, since there

[7] Community- level income could also capture a relative income effect. Conditional on individual-level income, a higher community-level income means that the individual is relatively poor, and a lower that she is relatively rich.

is more violence in general, which tends to increase female risk behavior, such as early sexual debut, as well as the number of rapes (Andersson et al., 2008).

Additionally, a relationship between inequality and HIV could exist because inequality is associated with more mobility, which seems to increase the spread of HIV (Oster, 2011). The most unequal societies in Sub-Saharan Africa tend to have an economic structure with large commercial farms and mines that generate geographical labor mobility. Since prostitution and transactional sex relationships are common in many of these places, migrant workers in Eastern and Southern Africa are more often infected than people in general, and there is a risk that they bring the disease to their home communities (GOM, 2004; Hargrove, 2008).

3. HIV/AIDS in Malawi

Malawi's first AIDS case was diagnosed in 1985, and from then on the epidemic spread rapidly, first in the major cities, and then in rural areas.[8] According to the most recent estimate, the national rate was 11% in 2009, which means Malawi registers the ninth highest HIV prevalence in the world (UNAIDS, 2010).

<<TABLE 1 ABOUT HERE>>

There are two main sources of information on HIV prevalence in Malawi, the 2004 MDHS and sentinel surveillance at antenatal clinics (ANCs). While the 2004 Malawi DHS are likely to provide good estimates of the prevalence rates in 2004, the ANC data is the only systematic information available of how the epidemic has evolved over time. UNAIDS uses the ANC data combined with MDHS data to estimate annual HIV rates, which are reported for selected years between 1990 and 2009 in Table 1. The prevalence rate rose from about 2% in 1990 to close to 14% at the end of the 1990s. During the 2000s, there was a decline to 11%, which indicates that at least prevalence is not increasing.

The relatively constant level of prevalence rate during the last 10 years hides very different geographical developments: the rates are declining in urban areas and increasing in rural areas. Urban HIV prevalence peaked at 26% in 1995 among women attending antenatal clinics, and then started to decline slowly. It was 17% in 2004. In the rural areas the prevalence rate reached 10.8% in 2004 (NSO

[8] See Arrehag et al. (2006) and Conroy et al. (2007) for more extensive descriptions of HIV/AIDS Malawi.

and OCR Macro, 2005; Republic of Malawi, 2006). There are also large differences across districts. Prevalence rates in some districts in Southern Region are as high as 20%–22%, with an average of 17.6%, while in Central and Northern Region they are on average 6.5% and 8.1%, respectively (National Statistical Office & ORC Macro, 2005). This is also reflected in prevalence among young women in our estimation sample; 13.6% in the South, 5.6% in the Center, and 6.4% in the North.

Furthermore, there are large age and gender specific differences. Table 1 show that HIV prevalence among women in the age group 15-19 is 9 times higher than for men, and 3.4 times higher in the age group 20-24.

In couples it is more common that only one of the two are HIV positive than that both are, as also seen in Table 1. It is more common that the man is the only HIV-positive partner, though the difference between men and women is not large.

Although Malawi's HIV epidemic is still unfolding, it seems to have reached a relatively mature stage. As evident from Table 1, national prevalence rates have not changed much during the last 10 years, and forecasts at the regional level indicate that the infection rates will remain stable the coming years (Geubbels and Bowie, 2007). Hence, the main drivers should have had time to affect the HIV rates across Malawi, making a cross-section analysis of a fundamentally dynamic process worthwhile.

4. Empirical model

To analyze the impact of economic inequality on HIV, we use a multilevel logistic model. It allows us to evaluate the effect of inequality at different levels on individual risk of HIV infection while accounting for other differences across communities, including unobserved ones. As opposed to aggregate level analysis, we can control for individual economic status, allowing for a non-linear effect on the probability of HIV infection. Thus we control for the effects of individual-level absolute poverty and wealth that could otherwise be confounded by inequality. Furthermore, we include measures of community-level economic status to control for possible society-wide effects of community poverty or wealth.

We introduce community effects at two different levels, the neighborhood, approximated by the sampling cluster, and the district. The probability of individual *i*, living in neighborhood *j* and district *d*, being HIV-infected is

$$
\begin{aligned}
&\Pr\left(HIV_i = 1\right) = logit^{-1}\left(ecstat_i \beta_{inc\_i} + x_i^I \beta_I + \alpha_{jd[i]}^{Neigh} + \alpha_{d[i]}^{Dist}\right) \\
&\alpha_{jd[i]}^{Neigh} \sim N\left(\beta_{inc\_n} ecstat\_n_{jd} + \beta_{ineq\_n} ineq\_n_{jd} + x_{jd}^I \beta_N, \sigma_{Neigh}^2\right) \\
&\alpha_{d[i]}^{Dist} \sim N\left(\beta_{inc\_d} ecstat\_d_d + \beta_{ineq\_d} ineq\_d_d + x_d^I \beta_D, \sigma_{Dist}^2\right).
\end{aligned} \tag{1}
$$

According to Eq. (1), the individual risk of being HIV infected depends on household wealth, $ecstat_i$, other individual-level characteristics, x_i^I, a neighborhood effect, $\alpha_{jd[i]}^{Neigh}$, and a district effect, $\alpha_{d[i]}^{Dist}$. The neighborhood and district effects depend on the economic status of a typical household and economic inequality, other community variables, and an unexplained part. The unexplained parts of the neighborhood and district effects are assumed to be normally distributed and independent of regressors.[9]

The assumption that the unexplained parts of the community effects are normally distributed is an improvement over assuming no community-level variation in addition to that captured by regressors, but the true variation might of course have a different distribution. As a robustness check, we therefore estimate models assuming a discrete distribution with a finite number of mass-points, where the probability that a unit belongs to a certain mass-point is estimated together with its locations.

Another potential concern is that the unexplained part of the community effect is assumed not to be correlated with the regressors. If we had used only individual-level regressors this assumption would certainly have been problematic; it is difficult to argue that individual poverty or wealth is not related to community characteristics that could matter for the spread of HIV. However, we assume individual-level poverty or wealth to be independent of community factors relevant for the spread of HIV *conditional* on community covariates, including the wealth of a typical household and economic inequality, a far less problematic assumption in our view. Still, as an additional check, we also estimate a model with fixed district effects, using district dummies.

Our dependent variable is HIV status. We know if an individual is HIV positive, but not when he or she was infected. If HIV-infected individuals who belong to certain groups survive longer than others, this could bias our parameter estimates. Thus, we restrict our sample to young women (age 15-24) who are

[9] The likelihood functions adherent to Eq. (1) is solved by numerical approximation using adaptive quadrature. More quadrature points gives better estimates but is more computationally demanding. To ensure that we use enough quadrature points we first estimated the model using 8 points and then 15 points. If the increase in quadrature points has no substantial effect on the log-likelihood value and the estimated parameters, we have enough quadrature points. A suggested rule of thumb is that the parameter should change by less than one percent.

likely to have been infected recently to make sure that our results are not influenced by differences in mortality. There are too few HIV-infected males in this age group to estimate the models, and including older men weakens the link to the neighborhood since many of them are mobile.[10]

5. Data and Variables

Our main source of data is the 2004 MDHS. This is the first nationally representative survey of HIV prevalence in Malawi, and the first to link HIV status with characteristics of the respondents and their household. There are 1,202 women aged 15-24 with available HIV status information. We also use data from the Integrated Income and Household Survey 1997/98 and the census from 1987 for measures of district-level median consumption, consumption inequality and population density, and data from the 2000 MDHS for measures of district mobility.

5.1 The HIV data and possible sample selection

In the 2004 MDHS sample, one third of the households were selected for HIV testing. The result of the test was not revealed to respondents.[11] As can be expected in a survey that collects information about sensitive issues, not all selected individuals could or wanted to participate, raising questions about the representativeness of the HIV-status sample.

There are two main groups with missing HIV status: respondents who were not interviewed, mainly due to absence, and respondents who were interviewed but refused to provide the blood sample for HIV testing. Out of the 1,665 selected and interviewed women aged 15-24, HIV status data was successfully collected for 72.2%.

In the final 2004 MDHS report, the issue of potential response bias is investigated by comparing observed and predicted HIV rates for different groups of people (NSO and ORC Macro, 2005). In general, observed and predicted rates differ little. The exception is Lilongwe District, where HIV status data was collected from only 39% of the selected women, and the observed HIV rate was 1.6%, while it was 15.1% in the rest of Malawi. An indicator of the size of the bias is that HIV prevalence among

[10] We also estimated models with men aged 15-29. The results for district inequality are very strong while the results for neighbourhood inequality are clearly weaker than among women age 15-24. These results are available from the authors on request.

[11] The data collection team were joined by a voluntarily testing and counselling (VCT) team that offered testing for those who were interested in knowing their HIV status.

women visiting antenatal clinics in Lilongwe District was not lower than in most other parts of the country (UNAIDS, 2004). Moreover, according to the 2004 MDHS report, the adjusted (estimated) prevalence rate in Lilongwe District was 11.5%. Thus, we exclude Lilongwe District from our analysis.[12] We also exclude the few observations from the small island Likoma, reducing the number of observations to 1,161 young women.

With an appropriate instrument, sample selection techniques could be used to correct for possible sample selection bias. Since we cannot think of any good instrument in our data we choose not to use sample selection techniques. What we can do, in addition to excluding observations from Lilongwe District, is to further investigate if consent to provide the blood sample is systematically related to known characteristics. We therefore estimated a probit model of the probability of consenting to the HIV test, which is reported in Table 2. Most importantly, economic inequality and household wealth are not related to the probability of consenting to provide blood for testing. Women provide the blood sample more often if they live in urban but poor neighborhoods and in rich districts with a mobile male population. Since infection rates are likely to be higher in such areas, the probability of HIV infection among young women in the sample could be biased upwards. However, we control for these area characteristics in all estimations. Furthermore, consent is not systematically related to marital status, having spoken about AIDS with the spouse, believing that people with AIDS are immoral, the number of non-spousal sex-partners, and never having had sex, which are all variables that are likely to be related to risk of HIV infection but are not included in our models.

<<TABLE 2 BOUT HERE>>

5.2 Explanatory variables

We measure our community variables at two different levels: the neighborhood, approximated by the sampling cluster (roughly a village), and the district. The major cities, Blantyre – the commercial center, Zomba – a university town in the South, and Mzuzu – 'the capital of the North', though formally part of larger districts, are treated as separate 'districts'. Lilongwe District, which includes the capital city Lilongwe, and Likoma District are excluded from the analysis as previously explained. In total we have 340 neighborhoods and 28 'districts'.

[12] As a robustness check we ran the main regression with the observations from Lilongwe, with no noticeable impact on estimated coefficients.

We measure individual-level economic status by the household wealth quintile, where wealth quintiles are based on a wealth index created using information on housing characteristics and a wide range of assets. The weights attached to each item in the index are the 'coefficients' of the first principal component in a principal components analysis. Similar wealth indices have been demonstrated to be good proxies for permanent income (Filmer and Pritchett, 2001). Neighborhood 'income' is measured by the wealth of the typical household, the cluster median of the household wealth index, and neighborhood inequality is measured by the household wealth index Gini coefficient.[13]

At the district level, 'income' is measured by the median level of consumption in 1997, and inequality is measured by the consumption Gini coefficient. The variables are from the Integrated Household Survey 1997-98 published in National Economic Council (2000) and NSO (2000), respectively.[14] One advantage with using data from well before 2004 is that the simultaneity problem is reduced since there cannot be feedback effects.

The literature on economic inequality and health, including the literature on economic inequality and HIV, focuses on income inequality, perhaps since data on income and income inequality tend to be easily available in rich countries. However, the theoretical links proposed in Section 2 are not obviously between income and health; consumption or wealth inequalities could be just as important. In fact, the relative income hypothesis and the society-wide effects hypothesis are about social distances and heterogeneity of the population, which might be better captured by expenditure or wealth.

Additionally, in agricultural-based societies such as Malawi consumption data tend to be favored over income data, since seasonal and rainfall variations cause large swings in income even within relatively small areas. Consumption also reflects wellbeing more directly than income (Deaton, 1997). Both consumption and wealth indices have been demonstrated to be good proxies of permanent income, and are more likely to be accurately measured than income. Moreover, our measure of wealth is constructed with observable items, which seems to be superior to measures that require valuation or recall information (Bollen et al., 2002).

[13] We also used the distance between the household wealth indices at the 90th and 10th percentiles, and a relative inequality measure proposed by McKenzie (2005) for asset indexes, as alternative neighbourhood inequality measures. The choice of measure does not have any impact on the results.

[14] Expenditure levels have been adjusted with 4 regional consumer price indices.

When used in levels, wealth indices, consumption, and income should in principle be relatively good representations of each other. However, wealth inequality is not necessarily a good representation of income inequality. This is because, even though the wealth index is likely to increase monotonically with income, the relationship is not likely to be linear (McKenzie, 2005). In principle, the findings of the empirical analysis are thus only valid for wealth, not income, inequality.

In our data, the measures of economic status and inequality are correlated with population density and closeness to urban areas. People in such areas are likely to be more mobile and interact with a larger number of people, which might increase the spread of HIV. In order not to confound this possible effect with wealth and inequality, we add a number of controls at both the neighborhood and the district level.

We use GPS coordinates of the sampling clusters to create measures of distances to road, to the closest of Malawi's four main cites, and to the most important border crossing to Mozambique (in the southeast along the main transport route). When computing the distance to road, consideration is taken to level curves, i.e. the distance around rather than across mountains is used. Distance to cities and the Mozambique border crossing is computed along roads and major paths. In DHS surveys that collect blood samples for HIV testing, a random error is added to GPS coordinates, creating measurement errors. This is, however, unlikely to lead to biases in our estimates, since the error is random. Finally, we have an indicator of urban residence at the neighborhood level.

At the district level we use population density in 1987 and mobility of the male population. Population density is calculated using data on district area and population from the Population and Housing Census in 1987. We have not been able to separate the three cities from their surrounding districts in creating the population density figures. The 2000 MDHS data set was used to create a district-level measure of the share of the district's male population that was mobile the previous year. A man is considered mobile if he was away throughout a whole month or on five or more different occasions during the past twelve months.

Finally, in the basic models we include dummies for the respondents' level of education: none or incomplete primary (reference category)', complete primary, and complete secondary or more; and age-dummies (15-19 (reference category), and 20-24. Education is likely to be related to income but may also capture attitudes as well as knowledge and ability to process information.

The risk of HIV infection might of course be related to a wide range of other factors, among them gender inequality, ethnicity, religion and male circumcision. However, we do not want to include more variables than necessary in our main estimations. Limiting the sample to only young women reduces it to 1,161 individuals, a fairly large number but most of these, 90%, are HIV negative. Still, as a robustness check we include individual-level indicators of all the above mentioned factors. We also try to investigate what might cause an association between inequality and HIV using indicators of sexual behavior, health, and migratory behavior as our dependent variables. Table A1 in the appendix provides variable definitions and summary statistics.

6. Results

6.1 Main estimations of the effect of inequality on risk of HIV infection

Results from the main estimations are reported in Table 3. Specification (1), our preferred model, is based on Eq. (1). In specifications (2) and (3) we relax the assumption that the unobserved part of the community effects is normally distributed, and approximate the distribution with discrete freely estimated mass-points: specification (2) has community effects at the neighborhood level and specification (3) at the district level.[15] We were not able to estimate the model with community effects at both the neighborhood and district levels; it did not converge. In specification (4) we use district dummies and normally distributed neighborhood effects.

To get a sense of the magnitude of the effects, we compute predicted probabilities of HIV infection for each individual in the sample under different scenarios. The predicted probabilities include the predicted unobserved effects First we set neighborhood economic inequality equal its mean less half a standard deviation, then we set it to its mean plus half a standard deviation. Comparing the predicted probabilities in these scenarios we get the effect of a one standard deviation increase in neighborhood inequality around its mean. The same procedure is repeated for district inequality, neighborhood median wealth, and district median consumption. We also compare predicted probabilities when household wealth is set to the poorest quintile, the second poorest quintile, the middle quintile, the second richest quintile, and the richest quintile. Table 4 reports the means of the predicted probabilities

[15] When estimating specification 2 and 3 we increased the number of mass-points by one until the likelihood did not increase, i.e. until the maximum Gateaux derivative was smaller than zero.

and Figures 1 to 5 show the cumulative distribution functions of the probabilities under the different scenarios. The predicted probabilities are based on the preferred model (specification 1).

<<TABLE 3 ABOUT HERE>>

As Table 3 reports, the effects of inequality are statistically significant at both the neighborhood and the district levels. This result is not altered when we estimate the distribution of the unexplained part of the community effects with discrete freely estimated mass-points (specification 2 and 3). The positive effect of neighborhood economic inequality also remains when we control for unobserved district factors with district dummies (specification 4).

An increase in either neighborhood (Figure 1) or district (Figure 2) economic inequality by one standard deviation around the mean creates a clear shift to the right (towards higher risk levels) in the cumulative distribution functions of the risk of HIV infection. The increases in neighborhood and district inequality raise the mean risk of HIV infection by 2.6 and 3.2 percentage points, respectively (Table 4).[16] Given a mean infection rate at about 10% for the women in our sample, these effects are sizeable.

Economic status of a typical household in the community does not have a consistent impact on the risk of HIV infection. When measuring it by median wealth in the neighborhood, there is no noticeable change in the risk of HIV infection as wealth increases by one standard deviation around the mean (Figure 3). The coefficient in Table 3 is not statistically significant in the main model, but positive and significant at the ten percent level in the semi-parametric model. However, when using median district consumption, living in a poorer district is associated with an increased risk of HIV infection (Figure 4 and Tables 3-4): the mean risk increases by 2.4 percentage points as district median consumption decreases by one standard deviation around its mean. In the main regression the coefficient is only statistically significant at the ten percent level, but in the estimation with semi-parametric district effects it is significant at the one percent level.

Young women from poorer households do not have higher risks of HIV infection (Table 3-4 and Figure 5). In fact, women from households in the middle and second richest wealth quintiles appear to have the largest risk of HIV infection, followed by women in the richest household wealth quintiles, while

[16] The increases in mean risk are 0.109-0.083 for neighbourhoods and 0.114-0.082 for districts.

women in the two poorest household wealth quintiles have the lowest risk. If all households belonged to the second richest wealth quintile (with the highest risk) rather than the second poorest one (with the lowest risk), the mean risk of HIV infection for women would increase by 4.3 percentage points (Table 4). However, the difference compared to the poorest group is only significant for the second richest in some specifications (Table 3). This seems to be due to our limited number of observations; the household wealth coefficients tend to have higher statistical significance in regressions with fewer covariates or when using a larger sample including older women (not reported).

Turning to the other control variables, women aged 20-24 have a higher risk of HIV infection than women 15-19. More education does not appear to be related to a different risk of HIV infection when household wealth is controlled for. Urban residence is associated with a higher risk of HIV infection, but this effect is not statistically significant when neighborhood distance measures are included. Living closer to the Mozambique border crossing along the main transport route in the southeast increases the risk of HIV infection, and, surprisingly, women who live closer to any of the four cities have a lower risk of HIV infection, but this is when we control for urban residence and other neighborhood distance measures.[17] We do not find any statistically significant effects of population density or mobility of the district's male population.

<<TABLE 4 ABOUT HERE>>

<<FIGURE 1 ABOUT HERE>>

<<FIGURE 2 ABOUT HERE>>

<<FIGURE 3 ABOUT HERE>>

<<FIGURE 4 ABOUT HERE>>

<<FIGURE 5 ABOUT HERE>>

[17] This result is reversed when the distance to the Mozambique border crossing is dropped.

6.2 Why is inequality associated with an increased risk of HIV infection?

In this section we first investigate whether the association between HIV infection and inequality can be related to differences in sexual behavior, general health, or return migration. Then we check if the results in Table 3 are robust to the inclusion of a number of other potential drivers of HIV in our model.

Table 5 reports multi-level regressions with sexual behavior indicators as dependent variables. Since young women's risk of HIV infection is not only affected by their own behavior, but also by that of their sexual partners and others in a common sexual network, we also consider men's and older women's sexual behavior. Reporting bias is likely to be a serious issue in survey data on sexual behavior, but we do not see any reason why it should be systemically related to inequality or wealth. The consequence should then be a classical measurement error problem with probable attenuation bias.

<<TABLE 5 ABOUT HERE>>

The first four specifications are multi-level ordered logistic estimations of the number of sexual partners other than the spouse/cohabitating partner during the last 12 months, for married women and men and for unmarried women and men. There are three possible categories; 0, 1, and 2 or more. District economic inequality increases men's reported number of sexual partners. The association is statistically significant at the five percent level among married as well as unmarried men; a one standard deviation increase in economic inequality increases the probability of reporting any non-spousal partner by 34.5% among married men and by 9.5% among unmarried men. There is also a positive and statistically significant association between unmarried men's reported sex partners and neighborhood inequality, but it is small. Unmarried, but not married, women report more non-spousal sex partners in more unequal places; a one standard deviation increase in neighborhood inequality increases the probability of reporting any sex partner by 4.5% and a one standard deviation increase in district inequality increases the probability by 5.8%, but this last association is only statistically significant at the ten percent level.

Specifications (5) and (6) are multilevel logistic estimations on abstinence and condom use at last non-spousal sexual encounter among young women. Abstinence in this case means never having had sex. District inequality, but not neighborhood inequality, is associated with a statistically significant smaller probability of abstinence, i.e. with an earlier sexual debut. The probability of abstinence decreases by 4.1% when district inequality increases by one standard deviation around its mean. Poverty also appears to be related to riskier sexual behavior since abstinence is less common among the poor

quintiles. Women in more unequal districts seem to use condoms more frequently, but the effect is not statistically significant.

Table 6 reports specifications with health indicators and return migration as the dependent variables in multilevel regressions. If inequality is associated with worse health, increased transmission rates among unhealthy populations could be one explanation for the impact of inequality on HIV. We use two indicators of general health, both closely related to undernourishment; anemia and stunting (children's height-for-age). These measures are general, but some specific health channels, such as malaria are related to anemia.[18]

Specification (1) is a multilevel logistic estimation of anemia among HIV-negative women. At later stages, HIV often leads to anemia, which is why we reduce the sample to only uninfected women. Specification 2 is a multilevel linear estimation of height-for-age, measured as the Z-score of children age 0-4. Rather than child characteristics, we include characteristics of the mother (age and level of education) in these specifications. Inequality is not associated with worse health when measured by anemia. When health is measured by stunting, inequality has a negative impact but it is not statistically significant. The effect of inequality on HIV does thus not seem to be mediated through health in general.

Migration could potentially cause both HIV and economic inequality. HIV prevalence is high among return-migrants, since many move to their home village when they fall ill with AIDS.[19] People returning from the city are also often wealthy compared to others in the village, and return migration could thus cause both the spread of HIV and inequality. The question is thus if this explains the association between HIV and economic inequality. Specifications (3) and (4) are multilevel logistic estimations of return migration for women and men in rural areas, where return migration is measured by a dummy variable equaling 1 if the respondent migrated from an urban to a rural area during the last five years. We find inequality to be associated with more return migration, especially among men: a one standard deviation increase in neighborhood inequality around its mean increases the probability of being a return migrant by 4.4% among women and 14.9% among men.

[18] We also included district–level malaria variables directly, based on incidence (Dzinjalamala, 2007) and ecology (Kazembe et al., 2006), but they had no impact. One explanation could be that they too crude measures.

[19] In the 2004 MDHS data, male HIV prevalence is higher among return migrants than among other men, and female HIV prevalence is higher among return migrants than among rural women or women who migrated to cities, but not higher than among urban women (not reported, but available from the authors).

<<TABLE 6 ABOUT HERE>>

Can temporary migration and links to cities explain the full effect of economic inequality on the risk of HIV infection among the young Malawian women? To evaluate this we add community return migration as an additional control (the share of return migrants both in the neighborhood and in the district) to specification 1 in Table 3, i.e. where we estimate risk of HIV infection. The impact of economic inequality only becomes somewhat weaker (Table 7, specification 1): at the district level, the inequality effect shrinks from 3.2 to 2.8 percentage points, but the coefficient is still statistically significant at the five percent level. At the neighborhood level the inequality effect decreases from 2.6 to 2.4 percentage points, while the coefficient now is significant at the five, instead of one, percent level.

<<TABLE 7 ABOUT HERE>>

Finally we control for various factors that have been suggested to matter for the spread of HIV. We add explanatory variables to specification 1 in Table 3, and report the results in Table 7. First, we control for religious affiliation and ethnicity (specification 2). Religions differ in terms of norms and traditions, and may matter for the spread of the epidemic. Cross-country studies regularly find that countries with a larger proportion of Muslims in the population have lower HIV rates (Sawers and Stillwaggon, 2010a). Ethnicity might also affect infection rates, most obviously since some cultural traditions involve sex (Malawi Human Rights Commission, 2006). Also, ethnicity, just as religion, may be related to norms and traditions that influence sexual behavior in general.

The inequality effects are barely affected by the inclusion of religion and ethnicity. However, religious affiliation seems to matter for the risk of HIV infection. Women belonging to the Presbyterian Church have a lower probability of HIV infection than catholic women. Muslims do not, in contrast what might be expected, have a lower risk of HIV infection. The effects of ethnicity on risk of HIV infection are generally not of importance when religion is controlled for.[20]

In specification (3) we add the share of men in the neighborhood who are circumcised. Male circumcision is not negatively related to HIV infection, as would have been expected from the findings of cross-country studies and controlled experiments (Gray et al., 2007). Indeed we find a higher risk of

[20] To save space, the religion and ethnicity dummy coefficients are not reported in Table 6.

HIV infection in areas with a higher prevalence of male circumcision, but including circumcision only reduces the inequality effects marginally. The increased HIV risk due to circumcision is probably related to unsafe practices. It is also possible that circumcision is practiced in groups with higher HIV risks due to other reasons.

Gender inequality is often considered an important driver of HIV (Gillespie, et al., 2007; Whiteside, 2008; Andersson et al., 2008), and may well be related to economic inequality. In specifications (4)-(6) we add gender inequality, differentiating between economic gender inequality, measured by women's participation in market work and the district gender gap in secondary schooling, and gender violence, which is measured by an indicator of whether the respondent's father ever beat the mother. Economic gender inequality does not appear to increase the risk of HIV infection, and it only affects the coefficient on inequality marginally. Gender violence, on the other hand, does increase the risk of HIV infection. It also weakens the inequality effects, especially at the district level, where it is reduced from 3.2 to 1.9 percentage points.

Lastly, in specification (7) we add interaction terms between community economic inequality and a dummy indicating that the woman belong to either the poorest or the second poorest wealth quintile. The purpose is to evaluate whether inequality increases the risk of HIV infection for all young women, or perhaps only for the relatively poor ones. The interaction terms are statistically insignificant; inequality appears to be bad for all women in the community. Again the pure inequality effects do not change much.

7. Conclusion

The aim of this study is to evaluate the impact of economic inequality on the spread of HIV/AIDS. We focus on Malawi, and analyze how inequality at both the neighborhood and district levels affects the individual-level risk of HIV infection among women aged 15-24. The analysis is carried out by estimating multilevel logistic models for individual women, which allow us to control both for unobserved community variation and estimate the impact of community-level explanatory variables. The main source of data is the nationally representative Malawi Demographic and Health Survey (MDHS) carried out in 2004, while district-level data was collected from various sources.

We find a strong association between economic inequality and the risk of HIV infection. Although a relationship between income inequality and HIV prevalence rates has been established at the cross-country level, as far as we know, this is the first study that shows a similar relationship using individual-level data for a particular country. When neighborhood-wealth inequality increases by one standard deviation around its mean, the risk of HIV infection for young Malawian women increases by 2.7 percentage points, and the effect of a similar increase in district-consumption inequality is to increase the risk of HIV infection by 3.4 percentage points. These effects are substantial, since mean levels of infection are about 10%.

So what might explain the inequality-HIV relationship? In all of our estimations we control for household wealth as well as individual education, allowing for non-linear relationships. Absolute poverty does not increase the risk of HIV infection for the women in our sample, and, since we control for it, poverty or diminishing health returns cannot explain the inequality HIV relationship in our study.

The effects of poverty are inconsistent: higher household and neighborhood median wealth seem to increase the risk of HIV infection, but the effects are statistically weak. On the other hand, lower median consumption at the district level is associated with higher risks of HIV infection; the effect is statistically significant only at the ten percent level in the main estimation model but at the one percent level in the estimation with semi-parametric district effects (and no neighborhood effects.

We find that economic inequality is associated with riskier sexual behavior. When district inequality is higher, men, both married and unmarried report more non-spousal sexual partners, and when neighborhood inequality is higher, unmarried women report more sexual partners. Furthermore, district inequality is associated with a lower probability of abstinence among young women, i.e. it is related to an earlier sexual debut.

According to economic theory, economic inequality ought to imply more transactional sex: in unequal places there are both more relatively poor women and more wealthy men that can afford transactional sex. If transactional sex is related to concurrent partners, the high level of transactional sex does not only constitute a high risk of HIV infection for the men and women who have transactional sex relationships, but for all in the sexual network. We have no transactional-sex-specific information in the data to test this hypothesis, but believe that more non-spousal sex partners and earlier sexual debut for young women are consistent with it.

Gender violence is often seen as a driver of HIV. Quite so, controlling for gender violence, measured by the share of men and women in the community who report that their father ever beat their mother, somewhat weakens the inequality impact. Moreover, gender violence has a clearly significant effect on the risk of HIV-infection.

We also find that migratory patterns can explain part of the inequality-HIV relationship. Economic inequality is related to an increased presence of return migrants in rural areas, and probably also to more out-migration to urban areas, temporary and permanent, increasing contacts with the cities. It is difficult to know exactly how causation runs in the migration-inequality-HIV relationship. Inequality could increase migration, but migration could also increase inequality and HIV as migrants bring back both wealth and the virus from the city to the village. When we control for the share of return migrants in our estimations of young women's HIV infection, the impact of neighborhood inequality is reduced by 10-15%.

We consider the possibility of an inequality impact on two measures of the general health situation; anemia among HIV negative women (aged 15-49) and height-for-age of children under age of 5. We do not find that inequality affects our health variables. These findings do not preclude that health in general, or specific diseases, are important drivers of the HIV epidemic in Malawi, only that our health indicators are not related to economic inequality.

When our findings are combined with those of other studies, there seems to be substantial evidence that economic inequality matters for the spread of HIV. This knowledge is of relevance both for the broader debate about economic inequality and development, and for the one about inequality and health. Moreover, many argue that it is necessary to address the underlying drivers of HIV risk, i.e., structural factors, to succeed with HIV prevention efforts (Rao Gupta et al., 2008), and economic inequality is an important structural factor. It is a challenge to reduce economic inequality, but in the medium term HIV prevention could focus on regions where economic inequality increases due to resource discovery or rapid structural change. In the short term, it seems sensible to concentrate on mediating factors such as sexual risk behavior and gender violence, addressing inequalities in gender relations and sexual interaction. Migration to cities is likely to be favorable for economic development, but knowledge about the risks it brings should be helpful for formation of HIV/AIDS prevention measures.

References

Abu-Raddad, L. J., Patnaik, P. & Kunlin, J. G. (2006). Dual Infection with HIV and Malaria fuels the Spread of both Diseases in Sub-Saharan Africa. *Science, 314*, 1603-1606.

Andersson, N., Cockcroft, A. & Shea, B. (2008). Gender-Based Violence and HIV: Relevance for HIV Prevention in Hyperendemic Countries in Southern Africa. *AIDS, 22*(Suppl.4), S73-S86.

Arrehag, L., de Vylder, S., Durevall, D. & Sjöblom, M. (2008). *The Impact of HIV/AIDS on Livelihoods, Poverty and the Economy of Malawi*. *Sida Studies No. 18*, Stockholm, Sweden: Sida.

Babones, S. J. (2008). Income Inequality and Population Health: Correlation and Causality. *Social Science & Medicine, 66*(7), 1614-1626.

Banerjee, A. & Somanathan R. (2007). The Political Economy of Public Goods: Some Evidence from India. *Journal of Development Economics, 82*(2), 287-314.

Barnett, T. & Whiteside, A. (2002). *AIDS in the Twenty-First Century: Disease and Globalization*. New Work, US: Palgrave Macmillan.

Bollen, K. A., Glanville, J. L. & Stecklov, G. (2002). Economic Status Proxies in Studies of Fertility in Developing Countries: Does the Measure Matter? *Population Studies, 56*(1), 81-96.

Bryceson, D. F & Fonseca, J. (2006). Risking Death for Survival: Peasant Responses to Hunger and HIV/AIDS in Malawi. *World Development, 34*(8), 1654-1666.

Bärnighausen, T., Hosegood, V., Timaeus, I. M. & Newell, M.-L. (2007). The Socioeconomic Determinants of HIV Incidence: Evidence from a Longitudinal, Population-Based Study in Rural South Africa. *AIDS, 21*(Suppl. 7), S29-S38.

Conroy, A. C., Blackie M. J., Whiteside, A., Malewezi, J. C. & Sachs, J. D. (2007). *Poverty, AIDS and Hunger: Breaking the Poverty Trap in Malawi.* New York, NY: Palgrave Macmillan.

Deaton, A. (2003). Health, Inequality, and Economic Development. *Journal of Economic Literature, 41*(1), 113-158.

Deaton, A. (1997). The Analysis of Household Surveys: A Microeconometric Approach to Development Policy. Baltimore, MD: Johns Hopkins University Press.

Dzinjalamala, F. (2007), Epidemiology of Malaria in Malawi. In E. Geubbels & C. Bowie (Eds.) *The Epidemiology of Malawi* (Chapter 3), Blantyre, Malawi: College of Medicine, University of Malawi.

Duesenberry, J. S. (1949). *Income, Savings, and the Theory of Consumer Behavior*. Cambridge MA: Harvard University Press.

Dunkle, K. L., Jewkes, R. K., Brown, H. C., Gray, G. E., McIntryre, J. A. & Harlow S. D. (2004).Transactional Sex Among Women in Soweto, South Africa: Prevalence, Risk Factors and Association with HIV Infection. *Social Science and Medicine, 59*(8), 1581-92.

Farmer, P. (1999). *Infections and inequalities: the modern plagues,* Berkley, CA: University of California Press.

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

Filmer, D. & Pritchett, L.H. (2001). Estimating Wealth Effects without Expenditure Data-or Tears: An Application to Educational Enrolments in States of India. *Demography 38*(1), 115-132.

Flemming, D. & Wasserheit, J. (1999). From Epidemiological Synergy to Public Health Policy and Practice: The Contribution of Other Sexually Transmitted Diseases to Sexual Transmission of HIV Infection. *Sexually Transmitted Infection*s, *75*, 3-17.

Fliessbach, K., Weber, B., Trautner, P., Dohmen, T., Sunde, U., Elger, C. E. & Falk, A. (2007). Social Comparison Affects Reward-Related Brain Activity in the Human Ventral Striatum. *Science, 318* (5854), 1305-1308.

Fortson, J. G. (2008). The Gradient in Sub-Saharan Africa: Socioeconomic Status and HIV/AIDS. *Demography*, *45*(2), 303–322.

Fox, A.M. (2010). The Social Determinants of HIV Serostatus in Sub-Saharan Africa: An Inverse Relationship Between Poverty and HIV? *Public Health Reports*, *125*(Suppl. 4), 16–24.

Geubbels, E. & Bowie, C. (2007). Epidemiology of HIV/AIDS in Adults in Malawi. In E. Geubbels & C. Bowie (Eds.). *The Epidemiology of Malawi* (Chapter 2). Blantyre, Malawi: Division of Community Health, College of Medicine, University of Malawi.

Gillespie, S., Kadiyala, S. & Greener, R. (2007). Is Poverty or Wealth Driving HIV transmission? *AIDS 21*(Suppl. 7), S5-S16.

Glymour, M. M. (2008). Sensitive Periods and First Difference Models: Integrating Etiologic Thinking Into Econometric Techniques: A Commentary on Clarkwest's 'Neo-Materialist Theory and the Temporal Relationship Between Income Inequality and Longevity Change'. *Social Science & Medicine*, *66*(9) 1895-1902.

GOM (2004) Behavioural Surveillance Survey (BSS) Report: Malawi 2004, Government of Malawi, Lilongwe, Malawi.

Gravelle, H., Wildman, J. & Sutton, S. (2002). Income, Income Inequality and Health: What Can We Learn from Aggregate Data? *Social Science and Medicine 54*(4), 577-589.

Gray, R. H., Kigozi, G., Serwadda, D., Makumbi, F., Watya, S. & Nalugoda, F. (2007). Male Circumcision for HIV Prevention in Men in Rakai, Uganda: a Randomized Trial. *The Lancet, 369(*9562), 657–666.

Halperin, D. T. & Epstein H. (2004). Concurrent Sexual Partnerships Help to Explain Africa's high HIV Prevalence: Implications for Prevention. *The Lancet, 364*(9428), 4-6.

Hargreaves, J. R., Bonell, C., Morison, L. A., Kim, J. C., Phetla, G., Porter, J., Watts, C. & Pronyk, P. M. (2007). Explaining continued high HIV Prevalence in South Africa: Socioeconomic Factors, HIV Incidence and Sexual Behaviour Change Among a Rural Cohort, 2001-2004. *AIDS 21*(Suppl. 7), S39-S48.

Hargrove, J. (2008). Migration, Mines and Mores: the HIV Epidemic in Southern Africa. *South African Journal of Science, 104*(1 & 2), 53-61.

Holmqvist, G. (2009). HIV and Income Inequality: If There is a Link, What Does it Tell us? Working Paper No. 54, International Policy Centre for Inclusive Growth, United Nations Development Programme.

Hunter, M. (2002). The Materiality of Everyday Sex: Thinking Beyond Prostitution. *African Studies, 61*(1), 99-120.

Jen, M. H., Jones, K., Johnston, R. (2009). Global Variations in Health: Evaluating Wilkinson's Income Inequality Hypothesis Using the World Values Survey. *Social Science & Medicine, 68*(4), 643-653.

Kazembe, L. N., Kleinschmidt, I., Holtz, T. H. & Sharp, B. L. (2006). Spatial Analysis and Mapping of Malaria Risk in Malawi Using Point-Referenced Prevalence of Infection Data. *International Journal of Health Geographic's, 5*(41).

Kjetland, E. Ndhlovu, P. D., Gomo, E., Mduluza, T., Midzi, N., Gwanzura, L., Mason, P. R., Sandvik, L., Friis, H. & Gundersen, S. G.(2006). Association between Genital Schistosomiasis and HIV in Rural Zimbabwean Women. *AIDS, 20*(4), 593-600.

Johansson-Stenman, O. & Martinsson, P. (2006). Honestly, Why are You Driving a BMW? *Journal of Economic Behavior and Organization.* 60(2), 129-146.

Lachaud, J.-P. (2007). HIV Prevalence and Poverty in Africa: Micro- and Macro-Econometric Evidences Applied to Burkina Faso. *Journal of Health Economics, 26*(3), 483-504.

Leigh, A., Jencks, C. & Smeeding, T.M. (2009). Health and Economic Inequality. In Salverda, W., Nolan, B. & Smeeding, T. M. (Eds.). *The Oxford Handbook of Economic Inequality* (384-405), Oxford, UK: Oxford University Press.

Lopman, B., Lewis, J., Nyamukapa, C., Mushati, P., Chandiwana, S. & Gregson, S. (2007). HIV Incidence and Poverty in Manicaland, Zimbabwe: Is HIV Becoming a Disease of the Poor? *AIDS, 21*(suppl. 7), 57-66.

Lurie, M. & Rosenthal, S. (2010). Concurrent Partnerships as a Driver of the HIV Epidemic in Sub-Saharan Africa? The Evidence is Limited, *AIDS and Behavior*, *14*, 17-24.

Luttmer, E. (2005). Neighbors as Negatives: Relative Earnings and Well-Being. *Quarterly Journal of Economics, 120*, 963-1002.

Mah, T. L. & Halperin, D. T. (2010). The Evidence for the Role of Concurrent Partnerships in Africa's HIV Epidemics: A Response to Lurie and Rosenthal. *AIDS and Behavior*, *14*(1), 25-28.

Mapingure, M. P., Msuya, S., Kurewa, N. E., Munjoma, M. W., Sam, N., Chirenje, M Z., Rusakaniko, S., Saugstad, L. F., de Vla, S. J. & Stray-Pedersen, B. (2010). Sexual behaviour does not reflect HIV-1 prevalence differences: a comparison study of Zimbabwe and Tanzania. *Journal of the International AIDS Society, 13*(45), doi:10.1186/1758-2652-13-45.

McKenzie, D. (2005). Measuring Inequality with Wealth Indices. *Journal of Population Economics, 18*(2), 229-260.

Mishra, V., Bignami-Van Assche, S., Greener, R., Vaessen, M., Hony, R., Ghys, P, Ties Boerma, J., Van Assche, A., Khan, S. & Rutstein S. (2007). HIV Infection Does not Disproportionately Affect the Poorer in Sub-Saharan Africa. *AIDS, 21*(Suppl. 7), s17-s28.

Mishra, V., Kottiri, B., Liu, L., Rathavuth, H., Khan, S. & Opio, A. (2008). The Association between Medical Injections and Prevalent HIV Infection: Evidence from a National Sero-Survey in Uganda. DHS Working Paper No. 42, Macro International Inc, Calverton, USA.

Morris, M. & Kretzscmar M. (1997). Concurrent Partnerships and the Spread of HIV. *AIDS 11*, 641-648.

Msisha, W. M., Kapiga, S. H., Earls, F. J. & Subramanian, S.V. (2008a). Socioeconomic Status and HIV Seroprevalence in Tanzania: A Counterintuitive Relationship. *International Journal of Epidemiology, 37,* 1297–303.

Msisha, W. M., Kapiga, S. H., Earls, F. J. & Subramanian, S.V. (2008b) Place Matters: Multilevel Investigation of HIV Distribution in Tanzania. *AIDS*, 22(6), 741-748.

National Economic Council (2000). Profile of Poverty in Malawi, 1998: Poverty Analysis of the Malawi Integrated Household Survey, 1997-98. Lilongwe, Malawi.

Nattrass, N. (2008). Sex, Poverty and HIV. CSSR Working Paper No. 220, University of Cape Town, South Africa.

NSO (2000). Malawi Integrated Household Survey 1997/98 - Statistical Abstract. Lilongwe, Malawi.

NSO (National Statistical Office of Malawi) and OCR Macro (2005). Malawi Demographic and Health Survey 2004 – Final Report. Zomba, Malawi, and Calverton, MD.

Oster, E. (2007). HIV and Sexual Behavior Change: Why not Africa? NBER Working Paper Number No. 13049.

Oster, E. (2011). Routes of Infection: Exports and HIV Incidence in Sub-Saharan Africa. *Journal of the European Economic Association,* forthcoming.

Over, M. (1998). The Effects of Societal Variables on Urban Rates of HIV Infection in Developing Countries: An Exploratory Analysis. In Ainsworth, M., Fransen, L. & Over, M. (Eds.), *Confronting AIDS: Evidence from the Developing World* (39-51), European Commission and World Bank, Brussels, Belgium, and Washington, DC.

Rao Gupta, G., Parkhurst J. O., Ogden J. A., Aggleton P. & Mahal A. (2008). Structural Approaches to HIV Prevention. *The Lancet, 372* (9640), 764-775.

Republic of Malawi, (2006) Report of the Malawi Triangulation Project: Synthesis of Data on Trends in the National and Local HIV Epidemics and the Reach and Intensity of Prevention Efforts. National Aids Commission, Lilongwe.

Sawers, L. & Stillwaggon, E. (2010a). Understanding the Southern African 'Anomaly': Poverty, Endemic Disease and HIV. *Development and Change*. *41*(2), 195-224.

Sawers, L. & Stillwaggon, E. (2010b). Concurrent Sexual Partnerships do not Explain the HIV Epidemics in Africa: a Systematic Review of Evidence. *Journal of International Aids Society*, *13*(34), 2-23.

Stillwaggon, E. (2006). *AIDS and the Ecology of Poverty*, Oxford, UK: Oxford University Press.

Stillwaggon, E. (2009). Complexity, Cofactors, and the Failure of AIDS Policy in Africa. *Journal of the International AIDS Society*, *12*, 12-20.

Subramanian, S.V. & Kawachi, I. (2004). Income Inequality and Health: What Have We Learned So Far? *Epidemiogical Review*, *26*, 78-91.

Swidler, A. & Watkins S. C. (2007). Ties of Dependence: AIDS and Transactional Sex in Rural Malawi. *Studies in Family Planning*, *38*(3), 147-62.

Tawfik, L. & Watkins, S. C. (2007). Sex in Geneva, Sex in Lilongwe, and Sex in Balaka. *Social Science and Medicine*, *64*, 1090-1101.

Tsafack Temah, C. (2009). What Drives HIV/AIDS Epidemic in Sub-Saharan Africa? *Revue d'économie du développemen*t, *23*(5), 41-70.

UNAIDS (2004). Epidemiological Fact Sheets on HIV/AIDS and Sexually Transmitted Diseases – Malawi 2004 Update. http://data.unaids.org/Publications/Fact-Sheets01/malawi_en.pdf.

UNAIDS, (2010). Global Report: UNAIDS Report on the global AIDS epidemic 2010. http://www.unaids.org/GlobalReport/Global_report.htm.

Veblen, T. (1899). The Theory of the Leisure Class - 2008 republication. www.forgottenbooks.org.

Wellings, K., Collumbien, M., Slaymaker, E., Singh, S., Hodges, Z., Patel, D. & Bajos N. (2006). Sexual Behaviour in Context: A Global Perspective. *The Lancet*, *368*, 1706-28.

Whiteside, A. (2002). Poverty and HIV/AIDS in Africa. *Third World Quarterly*, *23*(2), 313–332.

Whiteside, A. (2008). *HIV/AIDS: A Very Short Introduction*. Oxford, UK: Oxford University Press.

Wilkinson, R. G. & Pickett, K. E (2006). Income Inequality and Population Health: A Review and Explanation of the Evidence. *Social Science & Medicine*, *62*(7), 1768-1784.

Table 1: HIV prevalence rates among adults (aged 15-49) in Malawi

Estimated national prevalence rates 1990-2009							
	1990	1993	1996	1999	2002	2005	2009
	2.1	8.0	13.1	13.7	13.0	12.3	11.0
Prevalence rates in 2004 by gender and area							
	Urban	Rural	South	Central	North		
Women	18.0	12.5	19.8	6.6	10.4		
Men	16.3	8.8	15.1	6.4	5.4		
Total	17.1	10.8	17.6	6.5	8.1		
Prevalence rates in 2004 by gender and age-group							
	15-19	20-24	25-29	30-34	35-39	40-44	45-49
Women	3.7	13.2	15.2	18.1	17.0	17.9	13.3
Men	0.4	3.9	9.8	20.4	18.4	16.5	9.5
Prevalence rates in 2004 among couples by the woman's age							
	15-19	20-29	30-39	40-49			
Both are positive	3.1	7.1	9.4	4.1			
The man is positive	2.4	5.5	8.2	3.5			
The woman is positive	2.7	4.1	4.7	2.9			

Sources: UNAIDS (2008) and UNAIDS (2010) provide time series information on estimated national rates. The other information is from and NSO and ORC Macro (2005).

Table 2: Determinants of rejection to provide a blood-sample for HIV testing: Probit marginal effects

Age 20-24	0.00842
	(0.0887)
Second poorest	0.0291
	(0.124)
Middle wealth	0.107
	(0.127)
Second richest	0.0873
	(0.131)
Richest	0.0731
	(0.155)
Primary	-0.121
	(0.118)
Secondary	-0.233
	(0.157)
Urban	-0.405**
	(0.169)
Neighborhood median wealth	0.147**
	(0.0748)
Neighborhood inequality	-0.228
	(0.558)
Distance to road	0.00609**
	(0.00265)
Distance to city	-0.000474
	(0.000809)
Distance to border crossing	-0.000577**
	(0.000259)
District median consumption	-0.0650*
	(0.0364)
District inequality	0.0641
	(0.982)
District population density	0.000872
	(0.000925)
Mobility of district's male population	-2.142***
	(0.619)
Believe that people with AIDS are immoral	0.0196
	(0.0190)
Widowed or divorced	-0.00865
	(0.150)
Have never been married	0.0702
	(0.174)
Number of non-spousal partners last year	-0.173
	(0.158)
Never had sex	0.0702

	(0.174)
Constant	0.626
	(0.574)
Observations	1473

Clustered standard errors in parentheses.
*p<0.10, **p<0.05, ***p<0.01

Table 3: Main results of HIV infection among young women: Coefficients from multilevel logistic regressions

	(1) Neighborhood and district effects	(2) Semi-parametric neighborhood effects	(3) Semi-parametric district effects	(4) Neighborhood effects with district dummies
Individual- level regressors				
Age 20-24	1.816***	1.793***	1.782***	1.723***
	(0.303)	(0.298)	(0.293)	(0.283)
Second poorest	-0.0434	0.00113	-0.0608	0.0147
	(0.405)	(0.422)	(0.397)	(0.413)
Middle wealth	0.445	0.593	0.491	0.684*
	(0.373)	(0.379)	(0.366)	(0.370)
Second richest	0.539	0.787**	0.605	0.783**
	(0.378)	(0.380)	(0.371)	(0.380)
Richest	0.259	0.420	0.448	0.491
	(0.470)	(0.458)	(0.445)	(0.465)
Primary	-0.209	-0.295	-0.124	-0.134
	(0.354)	(0.344)	(0.341)	(0.354)
Secondary	0.0567	-0.122	-0.0398	0.165
	(0.440)	(0.424)	(0.434)	(0.444)
Urban	0.192	0.416	0.212	0.209
	(0.399)	(0.409)	(0.339)	(0.417)
Table 2 cont.				
Constant	-6.006***	-4.854***	-3.719***	-5.235***
	(1.598)	(1.623)	(1.402)	(1.131)
Neighborhood level regressors				
Median wealth	0.240	0.371*		0.157
	(0.203)	(0.211)		(0.217)
Inequality	4.494***	3.492**		3.211**
	(1.591)	(1.529)		(1.619)
Distance to road	-0.017	-0.013		-0.020
	(0.012)	(0.012)		(0.015)
Distance to city	0.007***	0.007**		0.005
	(0.002)	(0.003)		(0.005)
Distance to border crossing	-0.002***	-0.003***		-0.005**
	(0.001)	(0.001)		(0.003)
District-level regressors				
Median consumption	-0.201*		-0.265***	
	(0.109)		(0.101)	
Inequality	6.566**		6.090**	
	(2.711)		(2.720)	
Population density	-0.00406		-0.00279	
	(0.00266)		(0.00219)	
Male mobility	1.059		-0.596	

	(1.735)		(1.715)	
Unexplained community variance				
Cluster variance	0.115			0.000
	(0.380)			(0.000)
District variance	0.000			
	(0.000)			
Semi-parametric distribution				
Location 1st mass-point		-0.144	-2.172	
prob 1		0.975	0.122	
Location 2nd mass-point		1.929	0.301	
prob 2		0.019	0.878	
Location 3rd mass-point		16.123		
prob 3		0.007		
Observations	1097	1161	1097	1141
Log likelihood	-300.1	-330.2	-308.7	-303.0

Table 4: Means of predicted probabilities of HIV infection when we change the level of an explanatory variable

	Mean
Neighborhood inequality at its mean - 0.5 std. dev.	0.083
Neighborhood inequality at its mean + 0.5 std. dev.	0.109
District inequality at its mean - 0.5 std. dev.	0.082
District inequality at its mean + 0.5 std. dev.	0.114
Neighborhood median wealth at its mean - 0.5 std. dev.	0.089
Neighborhood median wealth at its mean + 0.5 std. dev.	0.103
District median consumption at its mean - 0.5 std. dev.	0.107
District median consumption at its mean + 0.5 std. dev.	0.084
Household wealth quintile=Poorest	0.079
Household wealth quintile=Second Poorest	0.076
Household wealth quintile=Middle	0.111
Household wealth quintile=Second richest	0.119
Household wealth quintile=Richest	0.097

Note: Predicted probabilities of HIV infection, for each individual in the sample, were computed based on Specification 1 in Table 2.

Table 5: Effect of inequality and income on sexual behavior – Multilevel regressions

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	Non-spouse partners	Non-spouse partners	Non-spouse partners	Non-spouse partners	Never had sex	Condom use non-spouse
Method	Ordered logit	Ordered logit	Ordered logit	Ordered logit	Logit	Logit
Sample	Married women	Married men	Unmarried women	Unmarried men	Young women	Young women
Second poorest	-0.54743	-0.057	-0.0174	0.115	-0.270*	-0.786
	0.39362	-0.355	-0.16	-0.251	(0.153)	(0.480)
Middle wealth	-0.45008	0.0276	0.0262	0.167	0.0558	0.331
	0.389279	-0.343	-0.157	-0.25	(0.148)	(0.399)
Second richest	-0.40576	0.202	-0.0142	-0.0051	0.273*	0.194
	0.400772	-0.347	-0.159	-0.241	(0.145)	(0.369)
Richest	-0.28254	0.19	-0.174	-0.0237	0.670***	0.583
	0.511867	-0.428	-0.183	-0.273	(0.171)	(0.418)
Neighborhood median wealth	0.013545	-0.0842	0.0906	0.0428	0.107	-0.134
	-0.33476	-0.212	-0.0773	-0.104	(0.091)	(0.155)
Neighborhood inequality	2.433188	0.385	2.026***	2.112**	-0.245	-0.94
	-2.00289	-1.36	-0.75	-1.077	(0.705)	(1.809)
District median consumption	-0.12222	-0.133	-0.072	0.0117	0.0238	0.018
	-0.1833	-0.0921	-0.0693	-0.0809	(0.066)	(0.112)
District inequality	-2.8327	5.296**	2.905*	4.979***	-3.309**	2.76
	-4.20961	-2.414	-1.56	-1.909	(1.629)	(2.785)
Observations	7380	1847	2841	979	4513	452
Log-likelihood	-360.6	-449.84615	-1531.3	-825.4	-1669.0	-243.9
	Effect of a one standard deviation increase in inequality around the mean (probability of a positive outcome or age in years)					
Neighborhood inequality	0.001 [8.5%]	0.002 [2.8%]	0.010 [4.8%]	0.001 [0.5%]	-0.001 [-0.4%]	-0.013 [-4.2%]
District inequality	-0.001 [-7.4%]	0.020 [34.4%]	0.012 [5.8%]	0.040 [9.5%]	-0.009 [-4.1%]	0.032 [10.9%]

All specifications also include controls for age, education and urban residence at the individual level, distance to road, city and main border crossing at the neighborhood level, and population density and mobility of the male population at the district level. They also control for unobserved neighborhood and district effects.
Standard errors in parentheses. Percentage changes in brackets.
*p<0.10, **p<0.05, ***p<0.01

Table 6: Effect of inequality and income on health and return migration – Multilevel regressions

	(1)	(2)	(3)	(4)
Dependent variable	Anemia	Height-for-age	Return migration	Return migration
Method	Logit	Linear regression	Logit	Logit
Sample	HIV negative women	Children age 0-4	Women in rural areas	Men in rural areas
Second poorest	0.0153	-52.73	0.0734	-0.280
	(0.150)	(94.81)	(0.163)	(0.381)
Middle wealth	-0.00463	-61.33	0.137	0.129
	(0.150)	(94.27)	(0.160)	(0.350)
Second richest	-0.0513	-55.60	0.275*	0.119
	(0.154)	(99.45)	(0.157)	(0.357)
Richest	-0.316	-129.6	0.552***	1.118***
	(0.201)	(130.7)	(0.177)	(0.394)
Neighborhood median wealth	0.0207	-9.672	0.833***	0.416
	(0.127)	(85.88)	(0.186)	(0.400)
Neighborhood inequality	0.721	-103.3	2.086***	3.225**
	(0.730)	(479.8)	(0.711)	(1.556)
District median consumption	0.0752	-21.83	-0.0341	-0.372***
	(0.0746)	(32.38)	(0.0851)	(0.116)
District inequality	0.0203	-345.1	1.546	4.550
	(1.838)	(828.5)	(1.991)	(2.841)
Observations	2001	7802	8996	2440
Log-likelihood	-1339.4	-72681.9	-1886.9	-476.1
Effect of a one standard deviation increase in inequality around the mean (probability of a positive outcome or Z-score)				
Neighborhood inequality	0.004 [1.1%]	-6.321 [1.1%]	0.003 [4.4%]	0.008 [14.9%]
District inequality	0.000 [0.0%]	-17.399 [2.9%]	0.002 [2.7%]	0.010 [17.0%]

All specifications also include controls for age, education and urban residence at the individual level, distance to road, city and main border crossing at the neighborhood level (except 3), and population density and mobility of the male population at the district level (except 4). They also control for unobserved neighborhood (except 3) and district effects (except 4). 2 and 3 use age and education of the mother rather than the children.
Standard errors in parentheses. Percentage changes in brackets
*p<0.10, **p<0.05, ***p<0.01

Table 7: The effect on the inequality health relationship from adding more explanatory variables.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Regression coefficients for the added variable(s) and neighborhood and district inequality							
Neighborhood return migration	1.791						
	(1.945)						
District return migration	-4.203						
	(5.351)						
Ethnicity dummies		Yes					
Religion dummies		Yes					
Neighborhood male circumcision			0.622*				
			(0.352)				
Women's market work (cluster)				-0.419			
				(0.575)			
Women's market work (district)				0.610			
				(1.108)			
Secondary education gender gap (district)					-0.155		
					(0.346)		
Father beat mother						0.725**	
						(0.299)	
Poor*neighborhood inequality							-0.281
							(0.420)
Poor*district inequality							2.415
							(4.848)
Neighborhood inequality	4.160**	4.718***	4.235***	4.744***	4.786***	4.963**	4.868***
	(0.167)	(1.566)	(1.531)	(1.631)	(1.684)	(2.102)	(1.725)
District inequality	5.681**	5.585**	5.367**	5.898**	5.948*	4.953	6.055**
	(2.834)	(2.791)	(2.698)	(2.886)	(3.132)	(3.586)	(2.944)
Effect of a one standard deviation increase in inequality around the mean (probability of HIV infection)							
Neighborhood inequality	0.024	0.028	0.026	0.027	0.027	0.023	0.028
District inequality	0.028	0.028	0.028	0.029	0.028	0.019	0.030

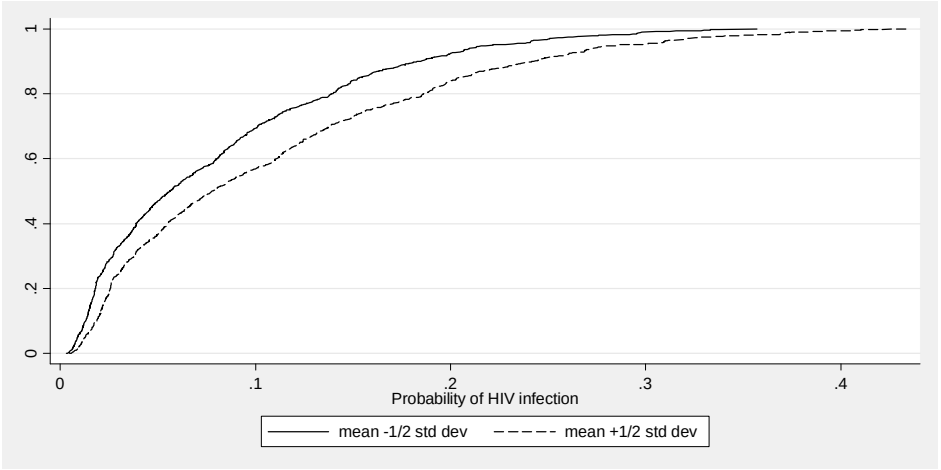
All specifications also include controls for age, household wealth, education and urban residence at the individual level, median wealth, distance to road, city and main border crossing at the neighborhood level, and median consumption, population density and mobility of the male population at the district level . They also control for unobserved neighborhood and district effects.
Standard errors in parentheses.
*p<0.10, **p<0.05, ***p<0.01

Table A1: Description of variables and summary statistics

Variable	Description	Obs.	Mean	Std. Dev.
	Young women with HIV-status information			
HIV-status	1 if HIV positive; 0 if HIV negative	1161	0.100	0.300
Age 15-19	1 if age 15-19; 0 else (reference group)	1161	0.439	0.497
Age 20-24	1 if age 20-24; 0 else	1161	0.561	0.497
Poorest	1 if poorest household wealth quintile; 0 else (reference group)	1161	0.180	0.384
Second poorest	Second poorest household wealth quintile; 0 else	1161	0.201	0.401
Middle wealth	Middle household wealth quintile; 0 else	1161	0.211	0.408
Second richest	Second richest household wealth quintile; 0 else	1161	0.215	0.411
Richest	Richest household wealth quintile; 0 else	1161	0.193	0.395
No education	1 if less than complete primary education; 0 else (reference group)	1161	0.094	0.292
Primary	1 if complete primary but not complete secondary education; 0 else	1161	0.703	0.457
Secondary	1 if complete secondary education; 0 else	1161	0.203	0.403
Urban	1 if urban residence; 0 if rural residence	1161	0.153	0.360
Chewa	1 if ethnicity is chewa; 0 else (reference group)	1161	0.294	0.456
Tumbuka	1 if ethnicity is tumbuka; 0 else	1161	0.126	0.332
Lomwe	1 if ethnicity is lomwe; 0 else	1161	0.171	0.376
Tonga	1 if ethnicity is tonga; 0 else	1161	0.022	0.148
Yao	1 if ethnicity is yao; 0 else	1161	0.163	0.369
Sena	1 if ethnicity is sena; 0 else	1161	0.034	0.182
Nkonde	1 if ethnicity is nkonde; 0 else	1161	0.008	0.088
Ngoni	1 if ethnicity is ngoni; 0 else	1161	0.102	0.303
Other ethnicity	1 if ethnicity is other than above; 0 else	1161	0.080	0.272
Catholic	1 if Catholic; 0 else (reference group)	1161	0.245	0.430
Ccap	1 if Central African Presbyterian Church; 0 else	1161	0.178	0.383
Anglican	1 if Anglican Church; 0 else	1161	0.022	0.145
Baptist	1 if Baptist/Seventh day Adventist; 0 else	1161	0.056	0.230
Other Christian	1 if other Christian church than above; 0 else	1161	0.351	0.478
Muslim	1 if muslim; 0 else	1161	0.145	0.352
No religion	1 if no religion/atheist; 0 else	1161	0.003	0.059
Father beat mother	1 if respondent report that her father ever beat her mother; 0 else	846	0.297	0.457
	Young women			
Never had sex	1 if the respondent never had sex; 0 else	4777	0.223	0.417
Condom with non-spouse	1 if the respondent used a condom last time she had sex with a non-spousal partner; 0 else	476	0.309	0.462
Non-spousal partners	Number of non-spousal sexual partners the last 12 months	4777	0.106	0.328
	Women			
Non-spousal partners	Number of non-spousal sexual partners the last 12 months	10776	0.073	0.333
	Table A1 cont.			
	HIV-negative women			
Anemia	1 if respondent has anemia; 0 else	2106	0.415	0.493

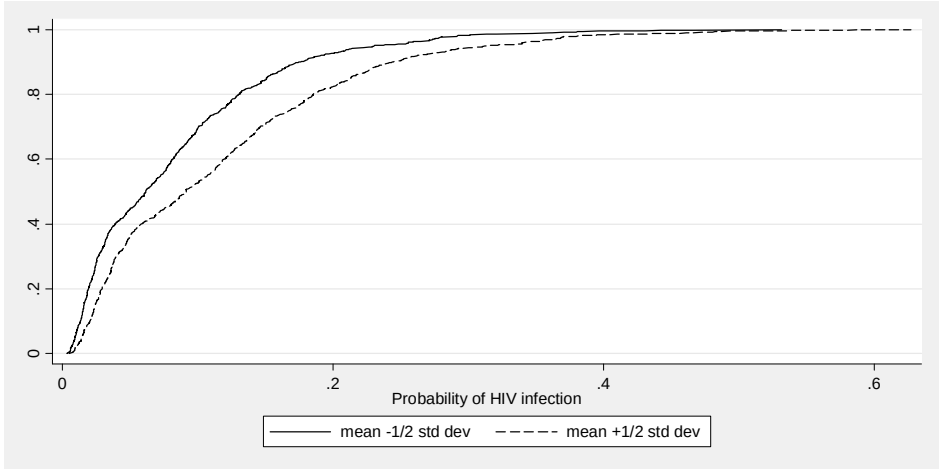
Women in rural areas				
Return migration	1 if respondent migrated from an urban to a rural area during the last five years	9370	0.063	0.242
Men				
Non-spousal partners	Number of non-spousal sexual partners the last 12 months	2953	0.243	0.698
Men in rural areas				
Return migration	1 if respondent migrated from an urban to a rural area during the last five years	2548	0.058	0.234
Young children (age 0-4)				
Height-for-age	Height-for-age Z-score multiplied by 100, where the Z-score is the child's height less the mean height for a child of that age in a reference population divided by the standard deviation of the mean height for a child that age in the reference population.	7802	-182.34	162.35
Neighborhoods. i.e. sampling clusters				
Median wealth	Cluster median of the household wealth index	484	0.875	0.738
Inequality	Household wealth index Gini coefficient	484	0.294	0.079
Distance to road	Distance in km to road	484	10.3	13.0
Distance to city	Distance in km to a city. the closest of Lilongwe. Blantyre. Zomba or Mzuzu	484	94.6	61.9
Distance to border crossing	Distance in km to the Mozambique border in the southeast along the main transport route	484	329.7	245.4
Return migration	Share in cluster that migrated from an urban to a rural area the last five years (men and women are weighted with 0.5 each)	484	0.054	0.067
Neighborhood male circumcision	Share of the men in the cluster that are circumcised	480	0.287	0.342
Female market work	Share of women in the cluster that participate in market work. i.e. that works for a money income	478	0.306	0.266
Districts				
Inequality	1997 Gini coefficient of household's per capita consumption in the district	26	0.412	0.064
Median consumption	Median of 1997 household's per capita consumption in the district	27	8.437	1.421
Male mobility	Share of the district's male population in 2000 that spent either at least one month away from home. or were away at least at 5 different occasions the last 12 months	28	0.330	0.084
Population density	District population density in 1987	28	112.0	83.6
Return migration	Share in district that migrated from an urban to a rural area the last five years (men and women are weighted with 0.5 each)	28	0.052	0.030
Female market work	Share of women in the district that participate in market work. i.e. that works for a money income	29	0.381	0.195
Secondary education gap	Share of women with secondary education/ share of men with secondary education	26	0.400	0.499

Figure 1: The effect of neighborhood inequality on the risk of HIV infection (Cumulative distribution functions of predicted probability of HIV infection).



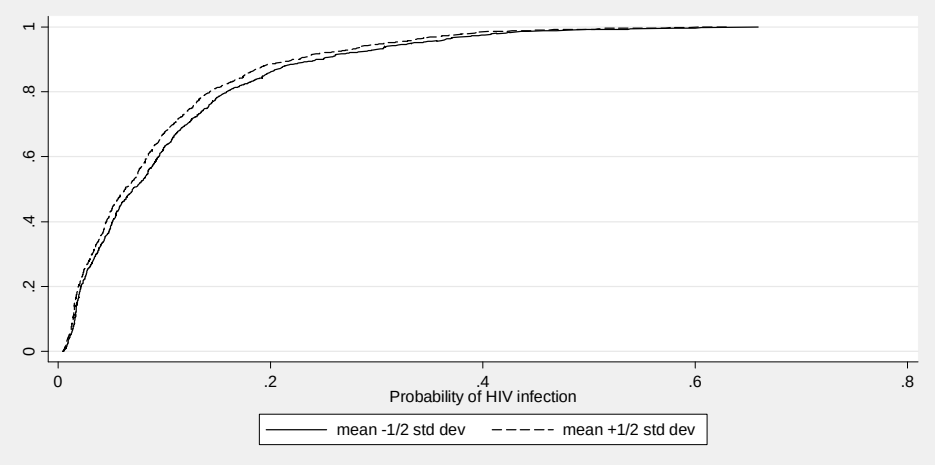
Note: Predicted probabilities of HIV infection, for each individual in the sample, were computed based on specification 3 in Table 2.

Figure 2: The effect of district inequality on the risk of HIV infection (Cumulative distribution functions of predicted probability of HIV infection).



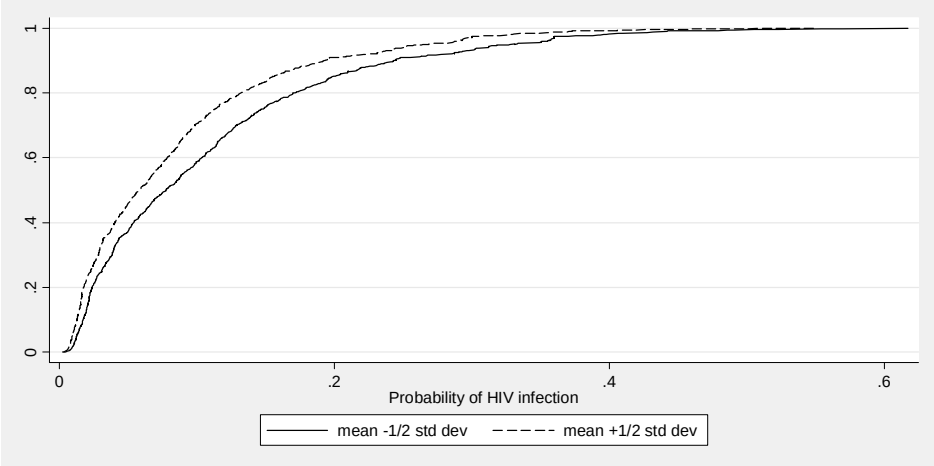
Note: See Figure 1.

Figure 3: The effect of neighborhood median wealth on the risk of HIV infection (Cumulative distribution functions of predicted probability of HIV infection).



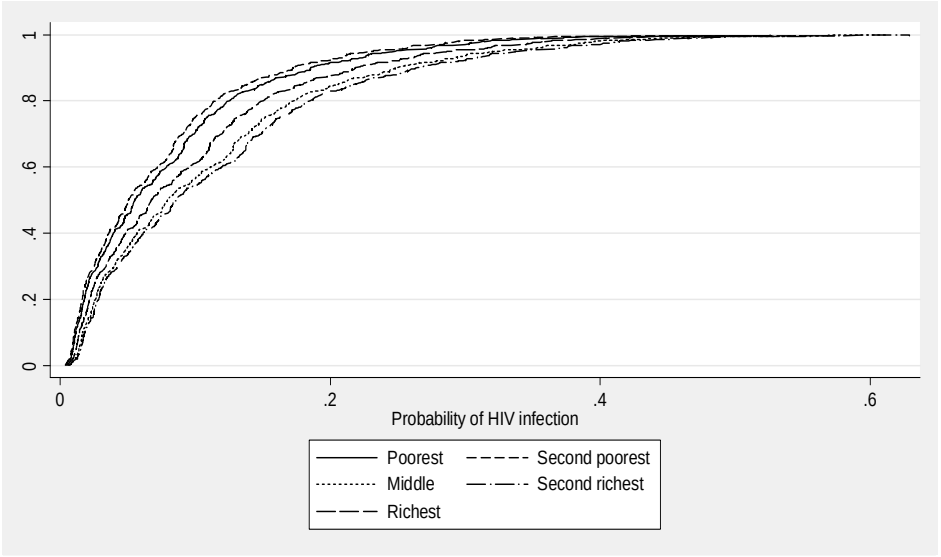
Note: See Figure 1.

Figure 4: The effect of district median consumption on the risk of HIV infection (Cumulative distribution functions of predicted probability of HIV infection).



Note: See Figure 1.

Figure 5: The effect of household wealth on the risk of HIV infection (Cumulative distribution functions of predicted probability of HIV infection).



Note: See Figure 1.