CSAE Working Paper WPS/2015-05

The Moral and Fiscal Implications of Anti-Retroviral Therapies for HIV in Africa¹

Paul Collier², Olivier Sterck³ and Richard Manning⁴

February, 2015

Abstract

Thanks to anti-retroviral therapies, people living with HIV in developing countries can now have a near-normal life at a cost of a few hundred dollars per year. We postulate that given this newly low cost of maintaining lives, there is a moral duty to rescue those who are infected. The core of the paper quantifies a reasonable lower bound for the fiscal consequences of this duty, which we show creates a financial quasi-liability which for some African countries is comparable to their debt-to-GDP ratios. Expenditures on prevention can pre-empt some of these liabilities. We construct a model to show that in some countries expenditure on prevention would be cost-effective, reducing liabilities by more than its cost. In principle, prevention should be pursued at least up to the point at which expenditure on it reduces the quasi-liability sufficiently to minimize the overall cost of accepting the duty to rescue. However, we show that even with optimal prevention the quasi-liability is likely to remain too high to be affordable for a significant number of African countries. Extending the model to two players, we show that if the international community accepts part of the quasi-liability, (as it does), it should finance an equal share of prevention and treatment efforts. Any imbalance in this distribution would introduce moral hazard and lead to a sub-optimal level of prevention.

¹ This paper was published as part of RethinkHIV. RethinkHIV is a consortium of senior researchers, funded by the RUSH Foundation, who evaluate new evidence related to the costs, benefits, effects, fiscal implications and developmental impacts of HIV interventions in sub-Saharan Africa. Other members of the consortium have made valuable contributions and suggestions on the paper, and we specifically acknowledge Rifat Atun, Angela Chang, Tim Hallett, Judith Kabajulizi, Mthuli Neube, Osondu Ogbooji, Michelle Remme, Sachin Silva, Mariana Siapka, Mikaela Smit, Anna Vassall, Charlotte Watts and Alan Whiteside. We also thank David Miller and Markus Haacker for useful comments and discussions.

² Blavatnik School of Government, and the Centre for the Study of African Economies, Oxford University, and the International Growth Centre.

³ Centre for the Study of African Economies, Department of Economics, Oxford University.

⁴ Blavatnik School of Government, Oxford University.
1. Introduction

Thanks to anti-retroviral therapies, people living with HIV in developing countries can now have a near-normal life at a cost of a few hundred dollars per year. Reflecting this perception that the medical battle against AIDS is now winnable, a cover story of *The Economist* was entitled ‘The End of AIDS’.\(^5\) That this should be the title of a journal focused on economics was, however, ironic. In this paper we show that as HIV is becoming less of a medical calamity, a corollary is that it is becoming a major new fiscal liability, and one that is currently insufficiently recognized.

The existence of affordable treatment has changed the ethics of treatment. In Section 2 we introduce and explore how the concept of the moral duty to rescue might apply to the provision of ART to Africans living with HIV (PLHIV). We are concerned to base this duty firmly within widely accepted moral norms rather than to advocate for some ethical ideal. In Section 3 we explore how the recognition of a moral duty to rescue has implications for the future financial liabilities of governments. This long-term liability can be estimated by establishing the present and future stocks of people who are HIV+ and have a claim to expenditures on ART. We show that for some African countries this quasi-debt is very substantial: the success of ART in reducing the threat to lives implies a large and growing threat to public finance. Section 4 makes the key step in the paper. While the moral duty to rescue requires expenditure on treatment, which can be conceptualized as a debt liability, expenditure on prevention can reduce this burden. For a state faced with long-term obligations arising from HIV, efficient resource allocation would therefore involve spending on prevention up to the point at which the marginal dollar reduces the discounted liability of treatment by at least one dollar. As such, prevention expenditures can be considered as an investment. In Section 5 we illustrate the optimal allocation of budgeting between treatment and prevention over time in the case of Malawi. Using circumcision as an example of a viable prevention strategy, we show that the extension of coverage would be a cost-effective way of reducing the fiscal burden of the moral duty to rescue. In Section 6 we consider the appropriate apportionment of long-term spending for HIV between African governments and the international community. How much of the financial burden created by the moral duty to rescue should be borne by high-prevalence countries? How much of this burden should be taken over by the international community? Who should bear the cost and the responsibility of investment in prevention?

2. The Moral Duty to Rescue and its Qualifications

Our starting point is the proposition that if with near certainty an early death can be averted and replaced by a healthy life at the cost of a few hundred dollars per person per year, most people would accept that at 21st Century global living standards there is a moral duty to rescue\(^6\), even though there may be disagreement as to where that duty lies. Supposing that the policies of democratic governments come to reflect the views of their citizens, our objective is to study the

---
\(^5\) June 2\(^{nd}\), 2011.
\(^6\) By using the terminology “moral duty to rescue”, we want to distance ourselves from the legal concept of the “duty to rescue” and the much-debated “rule of rescue” in ethics (Rulli, T., & Millum, J., 2014).
implications of this moral duty to rescue for the new economics of HIV. The moral duty to rescue people who are HIV positive need not imply that people recognize a broader obligation to help people poorer or less fortunate than themselves. While our analysis does not rest on anything special about HIV, the duty of rescue is independent of radical and demanding ethical frameworks such as Utilitarian Universalism. In Utilitarian ethics the overarching objective is the maximization of global utility and from this obligation other responsibilities could be derived for a myriad of redistributive actions. As argued by Jonathan Haidt (2012), there is clearly no global consensus for a Utilitarian Universalist ethics. In contrast, the duty to rescue is widely accepted as a moral responsibility even if wider responsibilities are denied. For example, there is a duty to pull a drowning child out of a public pond if it is within the power of a bystander to do so. Utilitarian Universalism might extend the duty of the bystander to building a fence around the pond to protect children from parental negligence. The important difference between the duty of rescue and such preventative actions is captured by the difference in the moral force of the statements ‘this child will die unless you do X’ and ‘there is a one-in-ten chance that each of ten unknown children will die unless you do X’. As Pinker shows in his analysis of how moral language is used, moral force is only generated by a specific line of responsibility between X and the certain effect on an identifiable person (Pinker, 2007).

The moral duty to rescue is not absolute. In a world of resource scarcity many desirable actions are not feasible, and so costs have to be taken into account alongside benefits. Applied to HIV, it is therefore conditional upon the cost of treatment relative to the capacity to bear that cost. In low-income societies in which a substantial percentage of the population is HIV positive, with many other calls on public revenues, citizens may not accept such a duty as reasonable. Privileging HIV sufferers would come at the opportunity cost of neglecting more pressing needs. However, we suggest that globally, a majority of the people above the world per capita average income of around $11,000 would find it morally unacceptable to abandon identifiable people to certain death rather than finance a $500 per year cost of rescue. In section 6, we will discuss how the financial resources needed for responding to this global moral responsibility should be apportioned between the states affected by HIV and the international community.

Nor is the moral duty to rescue independent of the behavioural choices of the person in jeopardy. A situation in which someone is in jeopardy because of persistent behaviour that a reasonable person would regard as foolhardily obviates a moral duty of rescue. Applied to health, a large body of evidence indeed finds that people give less priority to those who are considered to be responsible for their ill health (Dolan et al., 2005; Dolan, P., & Tsuchiya, A., 2009.). However, this in turn needs to be qualified. It is now well established that the decision-making capacity of people who face extreme

---

7 The International Covenant on Economic Social and Cultural Rights (1966), the 2006 and 2011 Political Declaration of the UN committing to moving towards universal access to HIV prevention, treatment and care, the Millennium Development Goal 6, the 90-90-90 target of UNAIDS, the actions of PEPFAR and the Global Fund to Fight AIDS, Tuberculosis and Malaria and the generalization of ART access in Africa are, among others, elements showing that the provision of ART to people living with HIV is widely regarded as a moral imperative by governments and their population. More generally, there is a global consensus that developed countries have a moral responsibility to work to reduce hunger, severe poverty and health problems affecting poor countries (Council on Foreign Relations, 2012).
poverty is impaired (Mullainathan and Shafir, 2013). For example, one key group of HIV+ people who have exposed themselves through risky behaviour is sex workers, but this choice is indeed often the result of desperate circumstances which mitigate responsibility for poor decisions. The moral duty of rescue generates a counterpart duty on the part of the person in need to respond. Thrown a rope, the reasonable person in the pond will seize it rather than insist upon a rescue requiring no effort. In the case of HIV, one moral duty to respond is to agree to be tested, as this is a necessary condition for the provision of ART. Government should therefore not only provide ART, but also organize stigma-free campaigns for people to test and get ART. Adherence, retention in care and engaging in safe behaviour are other examples of moral duty falling on the person rescued.

Some other health conditions also generate a moral duty of rescue. Left untreated they are certain to cause catastrophic consequences, while treatment is neither so expensive as to pose an unreasonable burden, nor so cheap that recipients can be expected to pay for it. Types-1 diabetes, tuberculosis and birth complications are examples of major health issues fulfilling these criteria. In this paper, we focus on HIV because two of its characteristics have profound economic and fiscal implications. First, an efficient response against HIV requires an optimal mix of prevention and treatment. Prevention can be valued as an investment to reduce the treatment cost of future infections. Furthermore, treatment itself has a preventative effect thereby generating positive spillovers in the long run. The complementarity between treatment and prevention will be analysed in section 4.

Second, while the treatments of many diseases require a short run, early intervention, ARV treatment must continue through the lifetime of the recipient. In this sense, ART is similar to the treatment of chronic illnesses. This long duration has two morally significant aspects. Firstly, once treatment has commenced the moral case for continuing it is far stronger than the case for starting it. The decision to start treatment rests, as we have argued, upon a moral duty to rescue: a specific action by A will prevent the death of B. The duty to continue treatment is stronger than this because to discontinue a program of treatment that was intended to be permanent would take a specific decision. That decision would be directly responsible for the death of a specific person. The decision to withhold treatment would require A to inflict certain and otherwise avoidable death on B: in other words it would meet the normal definition of murder. Treatment interruption has indeed been shown to increase the risk of disease progression and mortality (DART Trial Team DTT, 2008; Danel et al. 2006; El-Sadr et al., 2006). Hence, acceptance of the moral duty to rescue locks future decision takers into a commitment to continue treatment: it is a morally irreversible decision.

8 Even if sex workers are judged to have forfeited a moral duty of rescue, there is good reason to provide them with ART, and indeed specifically to target them. The early initiation of antiretroviral therapy has been shown to be a very efficient prevention tool, with the potential to reduce by as much as 96 percent the likelihood of transmission compared to delayed therapy (Cohen et al., 2011). As those who are engaging in high-risk behaviour are more likely to be infected, and to transmit the HIV virus, providing risk-taking individuals with early treatment is a cost-effective means of reducing the cost that would otherwise be incurred in rescuing those whom they might otherwise infect.

9 Studying the economic and fiscal implications associated with these other health conditions is important; it requires taking into account their specificities. This is beyond the scope of this paper.

10 This is also true for types-1 diabetes.
Further, once general provision of treatment has begun, it would be morally and politically odious to introduce restrictions upon access. Unlike other public services, the consequence of exclusion is highly discontinuous. People who had an expectation of continued life must be told that due to a change of policy they will be left to die whilst the service will continue to be provided to others.

The implication of irreversibility is that the decision to introduce treatment creates an on-going commitment that incurs a stream of future costs. Further, as we will show in the next section, the costs to the next generation may be significantly higher than the initial costs. It is therefore important that the decision be well informed as to these future costs. In this sense it is analogous to sovereign borrowing and the resulting long term obligations to service the debt (Haacker, 2011; Lule and Haacker, 2011): the people currently in charge of policy take on obligations that must be met by citizens far into the future. For this reason it is helpful to reformulate the ART decision in terms of the stock of liabilities rather than simply in terms of current expenditures. It is this feature that makes the application of the moral duty of rescue to HIV unique, and why in this paper we focus exclusively upon it. In the next section we undertake the task of converting flows into stocks.

3. From Flows of Costs to Stocks of Liabilities

The moral duty to rescue implies that the treatment of the population that is HIV positive generates a long-term financial liability. Country-by-country this is analogous to, and sometimes approximately commensurate with, the liability of sovereign debt. In this section we conceptualize and estimate this liability.

A set of eight sub-Saharan African countries were selected based on the sizes of their HIV infected populations and number of new infections. The selected countries accounted for about 45 per cent of the global HIV burden. The resources required for the moral duty to rescue were calculated assuming the following linear cost function:

\[ \text{Resources required}_{ijk} = \text{Reached population}_{ijk} \times \text{Unit cost}_{ijk} \]

where \( i=\text{country}, j=\text{year}, k=\text{service} \). In order to model the epidemic and derive estimates of the population benefiting from ART following the moral duty to rescue we used the SPECTRUM software (Futures Institute, 2014). SPECTRUM is a widely used modelling tool developed to support policy decisions concerning public health. It includes modules for HIV projections which are used, inter alia, by the UNAIDS secretariat to compiles estimates of HIV prevalence and incidence.

No ART program can realistically expect complete coverage of eligible people. UNAIDS has goals for the two components of coverage: diagnosis and treatment. Specifically, “90 per cent of all people living with HIV will know their HIV status” and “90 per cent of all people with diagnosed HIV infection will receive sustained antiretroviral therapy in 2020” (UNAIDS, 2014). These goals implicitly constitute a judgment as to what ART coverage the moral duty to rescue might reasonably
require, namely 81 per cent of the eligible population. Eligibility for ART is usually defined as a CD4 count threshold\(^1\). As shown on table 1, the median life expectancy of an HIV-positive adult between 25 and 34 years old who is not taking ART is expected to be 2.6 years when his/her CD4 drops below 200 cells/mm\(^3\). This rises to 4.4 years when the threshold is set below 350 cells/mm\(^3\) and 7.7 years when set below 500 cells/mm\(^3\). Which of these thresholds is sufficiently low and dramatic to generate a widely-shared moral duty to rescue is beyond the scope of the paper. The conventional approach adopted for such issues of medical ethics is the Bolam test, whereby if a standard is widely practiced by qualified professionals it is deemed to be reasonable.\(^12\) Currently, the 350 cell threshold probably best meets this condition. The more ambitious 500 cell threshold now encouraged by the WHO is not yet widely practised (Munthali et al. 2014), and an important part of its rationale is prevention as distinct from rescue.\(^13\) Prevention will be introduced in sections 4 and 5. In particular, we will discuss the financial impact of providing ART to individuals with CD4 count superior to 350 CD4/mm\(^3\) as a prevention strategy. In line with this, this section assumes that the coverage rates for other preventative interventions against HIV are equal to zero, except from circumcision rates which are assumed to remain constant.

<table>
<thead>
<tr>
<th>Age at infection</th>
<th>Median survival time without ART</th>
<th>Median survival time without ART</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;500</td>
<td>16</td>
<td>12</td>
</tr>
<tr>
<td>350-500</td>
<td>11.1</td>
<td>7.7</td>
</tr>
<tr>
<td>250-349</td>
<td>7.5</td>
<td>4.4</td>
</tr>
<tr>
<td>200-249</td>
<td>5</td>
<td>2.6</td>
</tr>
<tr>
<td>100-199</td>
<td>3.5</td>
<td>1.9</td>
</tr>
<tr>
<td>≤50-99</td>
<td>2</td>
<td>1.2</td>
</tr>
</tbody>
</table>

*Table 1 – CD4 count and median life expectancy (Spectrum estimates)*

Unit costs of treatment are assumed to be constant over time\(^14\) and are based on Schwartländer et al. (2011). Transition from 1st line to 2nd line ART was modelled assuming that 2.64% of people on

---

\(^{11}\) A useful indicator for measuring the progress of the HIV infection relates to the level of depletion of CD4 cells. These white blood cells are an essential part of the human immune system and can be infected and destroyed by the HIV virus. One mission of the World Health Organization (WHO) is to issue guidelines defining a CD4 count threshold below which ART is recommended. The new 2013 guidelines recommend that drugs be initiated as soon as a patient's CD4 count falls below 500 cells/mm\(^3\). Previously, the 2010 WHO recommendation was to offer treatment at 350 CD4 cells/mm\(^3\), and the 2006 guidelines set the threshold at 200 cells/mm\(^3\).

\(^{12}\) Segen’s Medical Dictionary. (2011). A more general procedural approach than the Bolam test for determining a consensual threshold is the “accountability for reasonableness framework” of Norman Daniels (Daniels, 2000; Daniels and Sabin, 2014). However, there is little basis for thinking that this framework would lead to a different conclusion.

\(^{13}\) In their report justifying the switch from the 350 CD4/mm\(^3\) to 500 CD4/mm\(^3\), the WHO write: “Expanding the eligibility criteria for ART and the options for using ARV drugs creates opportunities to save lives and reduce HIV transmission” (WHO, 2013)

\(^{14}\) The literature is unclear how the unit cost of providing ART is likely to change as coverage is expanded. For a given level of health facilities unit costs can be expected to decrease as ART coverage is expanded, thanks to economies of
treatment switch to 2nd line therapy every year (Renaud-Théry et al., 2011). Testing costs and health care system are omitted, implying that the estimates presented in this section are likely to underestimate the future cost of the moral duty to rescue. The detailed methodology used for producing estimates is explained in appendix A. 15

Table 2 presents different estimates of the cost of the duty to rescue for the selected eight countries. For the sake of comparison, the first three columns report the prevalence of HIV, the gross domestic product (GDP) per capita and the external debt as a percentage of the GNI of each country16. In column (4) we report the estimated cost in 2015 as a percentage of GDP, of full ART coverage defined on the CD4 count threshold of 350 cells/mm³. For example, in Malawi, the annual cost of ART provision is estimated at around 3.3 per cent of GDP17. These figures show that the cost of the duty to rescue is large relative to the GDP of the countries concerned, though of course it is negligible relative to global GDP.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Botswana</td>
<td>21.5 % adults</td>
<td>8,332</td>
<td>17.7 %GNI</td>
<td>1.04</td>
<td>18.9 %GDP</td>
<td>24.1 %GDP</td>
</tr>
<tr>
<td>Kenya</td>
<td>5.9</td>
<td>1,588</td>
<td>31.1 %GNI</td>
<td>0.34</td>
<td>N/A</td>
<td>9.1</td>
</tr>
<tr>
<td>Lesotho</td>
<td>22.7</td>
<td>1,390</td>
<td>31.3 %GNI</td>
<td>1.95</td>
<td>34.3 %GDP</td>
<td>73.6 %GDP</td>
</tr>
<tr>
<td>Malawi</td>
<td>8.6</td>
<td>275</td>
<td>31.7 %GNI</td>
<td>3.33</td>
<td>51.1 %GDP</td>
<td>80.3 %GDP</td>
</tr>
<tr>
<td>Nigeria</td>
<td>3.3</td>
<td>3,677</td>
<td>4.2 %GNI</td>
<td>0.15</td>
<td>3.4 %GDP</td>
<td>6.8 %GDP</td>
</tr>
<tr>
<td>South Africa</td>
<td>16.6</td>
<td>6,477</td>
<td>36.6 %GNI</td>
<td>0.63</td>
<td>13.2 %GDP</td>
<td>21.1 %GDP</td>
</tr>
<tr>
<td>Uganda</td>
<td>4.1</td>
<td>685</td>
<td>22.5 %GNI</td>
<td>0.72</td>
<td>11.9 %GDP</td>
<td>21.3 %GDP</td>
</tr>
<tr>
<td>Zimbabwe</td>
<td>13.5</td>
<td>1,073</td>
<td>75.5 %GNI</td>
<td>1.81</td>
<td>30.3 %GDP</td>
<td>38.7 %GDP</td>
</tr>
</tbody>
</table>

Table 2: Financial Implications of Long-term ART: the Moral duty to rescue as a Quasi-Debt

scale and scope (e.g. Siapka et al. 2014). However, typically existing facilities only attract around half of the population, so that expanding coverage beyond this to 81 per cent of the eligible population is likely to incur additional cost of providing facilities for the hard-to-reach. Further, these costs are likely to increase over time as income rises. Meyer-Rath and Over (2012), have shown that taking into account the need to expand health facilities implies a 42 per cent increase in the estimated cost of scaling up ART. The estimates we propose here are therefore likely to be lower bounds. 15 The limits of these estimates should be emphasized. The burden created the moral duty to rescue depends on parameters such as the future cost of ART, adherence rates, incidence and prevalence, testing rates, resistance to treatment, and the availability of a cure or a vaccine. Given the high degree of uncertainty surrounding these parameters, estimates of the total cost of the moral duty to rescue will have to be revised regularly.

16 GDP data comes from IMF projections. External debt data is based on World Bank indicators.

17 If a government tax revenues represent 15% of GDP and if 15% of this revenue is spent in health (Abuja target), the total amount invested in health by the government would represent 2.25% of its GDP.
Because antiretroviral therapies are life-long, the cost of treating the current generation of PLHIV will be recurrent. As shown in figure 1, for Malawi this cost is expected to increase for some years, not just absolutely but relative to GDP projected at the historic growth rate. Three factors drive this increase: some of those already infected but not yet needing treatment will cross the threshold of eligibility for ART; some people already on treatment will need more costly second-line therapies as their disease evolves; and new infections occur.

![Figure 1: cost of ART over time in per cent of GDP (GDP growth rate is assumed to be 4.3 per cent, which is the average growth rate of GDP between 1960 and 2013)](image)

Accepting a duty to rescue for the current generation of PLHIV therefore implies accepting a long-run flow of liabilities whose net present value can be compared to a debt. Columns (5) and (6) present the discounted present value of the cost of treatment to people currently HIV positive as a proportion of country’s GDP. The discount rate used for weighting the cost of future liabilities was set at 7 per cent which is an approximation of the marginal return on public investment in Africa (UNCTAD, 2013). This choice will be justified in the theoretical model that will be developed in section 4. By aggregating future liabilities discounted at this rate, we obtain the total amount that should be invested now to pay the future liabilities. Using such a high discount rate is conventional in policy-making (Gollier, 2011).

In column (5), the estimates do not include the cost of future infections. These estimates are therefore the barest bones of future liabilities, in which the only irreversibility allowed is to discontinue treatment for existing recipients. Even with this extreme restriction, it is apparent that

---

18 The timeframe assumed in GOALS/RNM is 2015 to 2050. In order to take into account the costs which are likely to occur after 2050, we do a terminal value calculation. In practice, we first divide the cost of the moral duty to rescue in 2050 by the discount rate to obtain the value of perpetuity. We then discount this value of perpetuity and add the amount to the total discounted cost of the moral duty to rescue.

19 While this may seem high, the rate of return on public investment in LICs measured by Canning and Bennathan (2000) is considerably higher than 10 per cent.
for some countries the duty to rescue would constitute a very significant addition to their sovereign liabilities. For Malawi and Lesotho, the addition represents more than 50 per cent of their GDP.

Columns (5) underestimate the future liability by considering only those people who are currently HIV positive. As shown in column (6), the financial liability generated by the moral duty to rescue is substantially greater when future infections are also taken into account. Furthermore, as shown in figure 2 for Malawi, a reduction in the discount rate would sharply increase the present value of the quasi-liability.

An implication is that the duty to rescue generates a new quasi-debt obligation that, for an unchanged rate of infection, would accumulate over the course of a few decades to become macro-significant. As with any such accumulation of quasi-debt, the issue of sustainability arises. Is there an ethically acceptable alternative to taking on this rising debt? Given the direct link between reducing the rate of infection and reducing the build-up of debt, the management of prevention is thus conceptually analogous to debt relief. Depending upon the scope for reducing the rate, it can potentially be of the same order of financial magnitude. We therefore turn from the estimation of the liability generated by the duty to rescue, to the optimal management of prevention.

4. Rescue and Prevention

Policy discussion of AIDS has already considered the balance of expenditures between treatment and prevention (e.g. Piot, P., Zewdie, D., & Türmen, T., 2002; Canning, 2006; Knight, R., Small, W., Pakula, B., Thomson, K., & Shoveller, J., 2014). To an extent this has reflected the tension between Utilitarian values, which favour prevention because it has the potential to save more people (albeit unidentified) in the future, and Prioritarian values, which usually privilege the treatment of the most
needy. In contrast, we show that far from these two choices being in tension, one implication of the moral duty to rescue is an optimal allocation between prevention and treatment.

We start from the proposition advanced in Section 2 that there exists a moral duty to rescue a large majority of those who are HIV positive. Since this duty to rescue incurs an obligation, those who bear it have a rational incentive to minimize the burden. Interventions that prevent a person becoming HIV positive are worth paying as long as their cost is less than the cost of the obligation. Since the implications of this proposition are the main focus of the paper, it is important to state it as clearly as possible. Money should be devoted to prevention at least up to the point at which the marginal dollar spent on prevention reduces the cost of the moral duty to rescue, as measured by the discounted cost of treatment, by one dollar.

In practice, the distinction between treatment and prevention is sometimes blurred because ART achieves both. Treatment as prevention (TasP) consists in initiating antiretroviral treatment earlier than what would be recommended on the basis of the medical needs of the patient, in order to protect partners from potential transmission. In serodiscordant couples, the provision of early antiretroviral therapies to HIV positive individuals has indeed the potential to reduce the likelihood of HIV transmission to the uninfected partner by 96 per cent compared to delayed therapy (Cohen et al. 2011). Antiretroviral drugs can also be provided to uninfected individuals as a pre-exposure prophylaxis (PreEP) to reduce HIV acquisition, with an efficacy in the range 44-75 per cent (Maartens et al., 2014). Further, they can be provided to pregnant and breastfeeding women for preventing mother to child transmission (PMTCT) (De Cock, 2000). Finally, they can be provided as a post exposure prophylaxis (PEP) shortly after an exposure to HIV to reduce the chances of infection. The preventative effect of ART implies that the provision of antiretroviral drugs does not have to be limited to those who have a claim on the moral duty to rescue. In particular, targeting PrEP and PMTCT towards individuals most likely to become infected and targeting TasP and PEP to those most likely to further transmit HIV may reduce future treatment liabilities so substantially that it fully justifies its cost. In line with our analysis, the provision of TasP, PrEP, PMTCT and PEP should be extended up to the point at which the marginal dollar spent on these prevention strategies reduces the cost of the moral duty to rescue by one dollar. In what follows, TasP, PrEP, PMTCT and PEP are categorized as prevention, as their provision is justified as an investment, but not directly as a moral duty. The preventative effect of antiretroviral drugs also justifies why it may be sensible to provide treatment to individuals who may be regarded by some as having forfeited their right to the moral duty to rescue. Because being on ART not only preserves their lives but also reduces the spread of infection to other people who are regarded as being more eligible for the moral duty to rescue, investing in their treatment is likely to be especially fruitful.

To investigate the interdependences between treatment and prevention further we use a simple two-period model. We consider a single country characterized by an HIV epidemic in the form of a two-period SI model. At time 0, a share $S_0$ of the population is uninfected, and a share $I_0$ is HIV infected. The population at time 0 is normalized to 1 such $I_0 = 1 - S_0$. HIV is a deadly disease and people at an advanced stage of infection are at risk of dying. Among those who are HIV positive, a
share $a_0$ is asymptomatic and do not need treatment to survive till period 1. The remainder, with share $(1 - a_0)$, needs ART to remain alive in period 1. The government determines the proportion $t_0$ of people in need who receive ART in period 0. According to the moral duty to rescue, $t_0$ should be equal to 1.

From period 0 to period 1, new infections occur. Incidence $N_0$ between the two periods depends first on the coverage of ART and second on the level of prevention in period 0, $p_0$. Without ART, the transmission rate of PLHIV is $b^h$. For people on ART, the transmission rate of HIV is denoted $b^l$. Because ART reduces the transmission rate of HIV (Cohen et al. 2011), we assume $b^h > b^l$. In our SI model, new infections are therefore equal to:

$$N_0 = (1 - p_0)[b^l(1 - a_0)t_0l_0 + b^h(a_0 l_0 + (1 - a_0)(1 - t_0)l_0)]S_0.$$ (1)

The first term in square brackets captures the infectivity of individuals on ART (low transmission rate $b^l$). The second term in square brackets represents the infectivity of asymptomatic individuals and individuals needing ART but not receiving it (high transmission rate $b^h$). A share $d$ of individuals needing treatment dies if they do not get ART.

In period 1, the number of uninfected individuals is $S_1 = S_0 - N_0$. The number of infected individuals is given by $I_1 = I_0 + N_0 - I_0 (1 - a_0)(1 - t_0)d$. The number of people needing treatment in period 1 is equal to the number of people who needed treatment in period 0 and who survived, plus a proportion $(1 - a_1)$ of people who are infected in period 1 and who did not need ART in period 0.

The complementarity between prevention and treatment can be illustrated within this framework. Let us assume that the government recognizes the moral duty to rescue and therefore imposes full coverage of ART, that is, $t_0 = t_1 = 1$. If the unit cost of ART is equal to $\pi^{20}$, the total cost of the moral duty to rescue will be equal to $(1 - a_0)l_0 \pi^t$ in period 0 and $[(1 - a_0)l_0 + (1 - a_1)(a_0 l_0 + N_0)] \pi^t$ in period 1. As incidence $N_0$ is a decreasing function of prevention, $p_0$, prevention can be used to diminish the total cost of the moral duty to rescue. Only the cost of future infection is affected by prevention. It is therefore helpful to conceptualize the total cost of the rule of rescue as having a fixed part, which depends on past infections, and a variable part which arises in the future and which can be optimized thanks to prevention. As shown in Table 2, this latter part represents 25-50 per cent of the total cost induced by the moral duty to rescue.

In order to optimize prevention, its cost has to be characterized. It is constituted of two elements. First, the implementation of prevention programmes has a direct cost. Because prevention methods
vary in cost-effectiveness and the costs of reaching different populations varies, the cost of prevention is expected to be an increasing and convex function $c(p_0)$ of prevention.

Second, money invested in prevention has an opportunity cost. If the government is not budget constrained in period 0, the opportunity cost of investing in prevention is determined by the marginal rate of return on alternative investments. If, however, the government has to borrow to finance prevention efforts in period 0, it faces the cost of debt servicing in period 1. We assume that the marginal interest rate on investment is equal to the cost of borrowing and denote this rate $r$.

In order to optimize prevention, the government has to minimize the following objective function:

$$\min_{p_0} [(1 - a_0) l_0 + (1 - a_1) (a_0 l_0 + N_0)] \pi^t + (1 + r)c(p_0)$$

(2)

where $N_0$ is given by equation (1). The associated first-order condition can be expressed as follows:

$$\frac{(1-a_1) [ a_0 (b^h-b^i) + b^i ] l_0 (1-l_0) \pi^t}{(1+r)} = c'(p_0)$$

(3)

The left-hand side of this equation represents the marginal discounted cost of ART treatment. The right-hand side is the marginal cost of prevention. The two sides of this equation and its solution are represented in figure 3.

![Figure 3: Optimizing prevention to reduce the financial burden created by the moral duty to rescue.](image)

Formally, solving the first-order condition gives the optimal level of prevention:
\[ p_0 = c^{-1} \left( \frac{(1-a_1) [a_0 b^h + (1-a_0) b^s] l_0 (1-l_0) \pi^0}{(1+r)} \right) \] (4)

Five propositions can be derived from the comparative statics of \( p_0^* \).

1. \( \frac{\delta p_0^*}{\delta c^t} > 0, \frac{\delta p_0^*}{\delta \pi^t} > 0 \): The optimal level of prevention is decreasing in its marginal cost and is increasing in the cost of treatment. The higher the cost of prevention, the lower the attractiveness of investment in prevention. The higher the future cost, the higher the incentive to invest in prevention.

2. \( \frac{\delta p_0^*}{\delta r} < 0 \): Prevention is decreasing in the marginal interest rate. If the marginal interest rate is high, the opportunity cost of prevention is high, and the optimal prevention level will be correspondingly lower.

3. \( \frac{\delta p_0^*}{\delta b^h} > 0, \frac{\delta p_0^*}{\delta b^s} > 0 \): The optimal level of prevention is increasing in the transmission rate of HIV, both with and without ART. Investing in prevention is more important if infectivity of HIV is high.

4. \( \frac{\delta p_0^*}{\delta a_0} > 0 \): The optimal prevention level is decreasing in the proportion of PLHIV who need treatment in period 0. If HIV is considered in a dynamical framework, the proportion of PLHIV needing treatment in period 0 is expected to be low, either in the early stage of the increase in ART coverage, or if the CD4 count threshold determining who has access to free ART is low. In this case, many PLHIV initially have too high a CD4 count to be eligible for ART. As their infectivity is expected to be high without ART, high investment in prevention will be needed. Conversely, if the proportion of PLHIV needing treatment in period 0 is high, the average infectivity of PLHIV will be low and investment in prevention should be lower.

5. \( \frac{\delta p_0^*}{\delta a_1} < 0 \): Prevention is increasing in the proportion of people newly infected who need ART. If PLHIV need ART early on, the cost the moral rule of rescue will be increased, thereby increasing the attractiveness of prevention. This will be the case if the CD4 count threshold that determines eligibility for ART is set high.

Taken together, Propositions 4 and 5 imply that the relationship between optimal prevention and the CD4 count threshold determining eligibility is ambiguous and so can only be determined empirically.

Given the reduced transmission rate of HIV while on ART, (Cohen et al. 2011), the increase in the coverage of ART implied by acceptance of the moral duty to rescue is likely to put the trajectory of the HIV disease on a decreasing trend. Within our theoretical framework, accepting the moral duty to rescue is equivalent to imposing \( t_0 = t_1 = 1 \). This has two opposite effects on the number of PLHIV. Mortality from HIV is lowered, thereby increasing the number of PLHIV. However, the reduced infectivity of PLHIV who are on ART is expected to reduce incidence and hence reduce the number of PLHIV. This latter effect is expected to be more important than the former in the long run, especially if prevention is introduced as determined in equation (4). This is illustrated in figure 4 for Malawi. Without either ART or investment in prevention, the number of people living with HIV.
in Malawi is expected to increase far into the future, because of high infection rates and high population growth. The number of people living with HIV can be lowered in the long run compared to this baseline scenario if ART coverage is increased. In the short run the number of people living with HIV would be a little higher with full ART coverage as treatment reduces mortality. However, introducing prevention in the form of circumcision would rapidly and substantially reduce the number of people living with HIV.

As explained in section 2, the moral duty to rescue is an undemanding moral framework, and so is likely to be widely accepted by the population. In contrast, the moral framework of Universalist Utilitarianism, often used in health economics, is radically more demanding and so a democratically empowered electorate may not accept the myriad of redistributive actions that would be desirable according to it. Hence, actions warranted according to the moral duty to rescue framework set a lower bound for guiding the response against HIV. The warranted actions which we have derived also constitute a lower bound because we have assumed that those kept alive by treatment suffer no adverse effects from HIV. In reality, HIV has a wide impact than on the immune system, affects mental health, and may induce stigma and fear. Further, the side-effects of treatment worsen with age. HIV is now a long-term chronic disease for many patients so that the HIV-patient population is ageing. Older patients are increasingly suffering from age-related non-communicable diseases which may be associated with HIV-infection and treatment. Numerous studies have found an increased risk for many NCDs in HIV-patients compared to age-matched uninfected population.”

21 Goulet et al. 2007; Guaraldi et al. 2011; M. S. Freiberg et al. 2013; Schouten et al. 2014
including cardiovascular disease\textsuperscript{22}, non-AIDS malignancies\textsuperscript{23}, liver and kidney disease\textsuperscript{24}, and osteoporosis\textsuperscript{25}. The additional costs of these sides-effects and co-morbidities are not incorporated in the estimates used in this paper because they may not meet the criteria that trigger the moral duty to rescue. Nevertheless, these further potential costs underline the significance of HIV-infection as a costly life-long condition. Governments wanting to be more generous than the moral duty to rescue framework could take these adverse consequences into account. According to proposition 1, this would increase the optimal investment in prevention. Our estimates thereby provide a lower bound for the liabilities that governments will face in terms of prevention and treatment.

In this section, we showed that, analogous to any other inter-temporal optimization problem, investment in prevention should be increased up to the point at which its marginal cost equals the marginal reduction in the discounted total liability created by the moral duty to rescue. Characterizing the two sides of this equation so as to derive optimal expenditures on particular prevention strategies is likely to be an empirical challenge given the complexity of data required. The determination of the optimal prevention level will be illustrated in next section using data for Malawi. In section 6, we extend this model to a two-country framework to discuss how investment in prevention should be apportioned between high prevalence countries and the international community.

5. Optimal Investment in circumcision the Case of Malawi

In section 4, we showed that prevention can be understood as an investment in the reduction of future liabilities brought about by the moral duty to rescue. We derived an optimal decision rule which states that prevention should be increased up to the point at which the marginal dollar spent in prevention equals the marginal cost of the duty to rescue. In practice, the cost of the duty to rescue is determined by the annual cost of ART cumulated over time and discounted over the expected duration of treatment. Matching this cost to the marginal benefits of prevention is likely to be empirically complex. For each possible intervention that prevents HIV infection, a schedule of the marginal efficacy of expenditure must be estimated. Such analysis must be context-specific as there is no global, or even African, marginal cost of treatment, or marginal efficacy of prevention (Anderson et al. 2014). As several prevention interventions with reinforcing benefits exist in each context, an optimal mix of interventions need to be considered when selecting prevention interventions. This analysis also needs to go beyond the health sector as narrowly conceived, as many effective prevention strategies are multi-sectorial in nature (Remme et al., 2012, Remme et al., 2014). To our knowledge, there is as yet no epidemiological model sufficiently sophisticated for such an optimization.

\textsuperscript{22} Klein et al. 2002; Triant et al. 2007; Currier et al. 2008; M. Freiberg et al. 2011; Cole et al. 2004
\textsuperscript{23} Kirk et al. 2007
\textsuperscript{24} Neuhaus et al. 2010; Odden et al. 2007
\textsuperscript{25} Triant et al. 2008; Brown and Qaqish 2006
While it is beyond the scope of the present analysis to build such a model, we illustrate what its potential might be by focusing on two prevention strategies, treatment-as-prevention and circumcision. For a particular context, Malawi, we use SPECTRUM to illustrate how these prevention strategies could be used to minimize the cost of the duty to rescue.

Given the large preventative effect of providing early antiretroviral therapies to HIV positive individuals (Cohen, 2011), treatment-as-prevention proposes to extend the CD4 count threshold giving access to ART. In reality, the proposed increase in the threshold from 350 to 500 has been justified partly as a prevention strategy, and partly as a way of saving lives, and hence rescue. But for purposes of illustrating the implications for prevention that follow from the duty to rescue, we suppose that the latter only warrants the 350 threshold, so that the difference between that and any higher threshold constitutes prevention. The issue then becomes whether such a prevention strategy sufficiently reduces future infections that it reduces the overall burden of the duty of rescue. In investigating this issue, as previously we used SPECTRUM, estimating the total cost of ART provision for different CD4 count thresholds. As shown in Figure 5, we find that ART as prevention is not sufficiently effective to reduce the quasi-liability generated by the moral duty to rescue. In this it contrasts with most existing studies on the cost-effectiveness of earlier eligibility for ART which usually conclude that treatment-as-prevention is a very cost-effective prevention strategy (Eaton et al., 2014, Walensky et al, 2012). While our analysis is meant to be illustrative rather than definitive, it is worth noting why our results are so different. First, the discount rate used in the global health literature is usually lower, set at 3 per cent, implying that current investments in prevention are more valued than in our framework. As we discuss in the next section, such a low discount rate may be a reasonable assumption for a donor, but would be hard to justify if the finance is provided by an African government. Second, in other studies the criterion is ‘cost-effectiveness’. For example, Walensky et al. (2012) and Eaton et al. 2014) classified early ART as cost-effective if its incremental cost per DALY averted or per life-year saved was less than three times annual per capita GDP. In comparison, the moral framework introduced in this paper requires prevention to be cost-saving: the net present value of investments in prevention should be lower than the net present value of the ensuing reduction in the cost of the moral duty to rescue. In other words, the value of an infection averted is given by the cost of ART provision. This cost is expected to be much lower than GDP per capita, especially in middle-income countries.
Circumcision has been shown to reduce by 60 per cent the risk of acquiring an HIV infection for men (Auvert et al. 2005; Bailey et al., 2007). We show that the total cost of the moral duty to rescue can be considerably reduced by investing in circumcision. To reach this conclusion, we start from the ‘no prevention’ scenario of section 3 and increase circumcision rates step by step in order to assess how circumcision can be used as an investment strategy to reduce future liabilities.

According to our estimates, increasing circumcision coverage is a very cost effective prevention strategy for Malawi (Table 3). This strategy would require making a considerable initial investment to achieve a quantum increase in the coverage of circumcision. This initial investment is worth doing because the reduced incidence and prevalence that it generates would significantly diminish the subsequent yearly cost of the duty to rescue. An investment of $267 million in 2015 to increase circumcision coverage would reduce the net present value of the cost of ART provision by $734 million. The net pay-off of this investment would be worth $468 million; the total liability brought about by the duty to rescue – i.e. the net present value of ART and circumcision costs - would be reduced from $4.02bn to $3.55bn.
Table 3 – Estimates of the total cost of treatment and prevention for different circumcision coverage

<table>
<thead>
<tr>
<th>Coverage</th>
<th>Total Cost of Treatment</th>
<th>Total Cost of Prevention</th>
</tr>
</thead>
<tbody>
<tr>
<td>60%</td>
<td>3,774</td>
<td>75.7</td>
</tr>
<tr>
<td>70%</td>
<td>3,714</td>
<td>74.5</td>
</tr>
<tr>
<td>80%</td>
<td>3,657</td>
<td>73.4</td>
</tr>
<tr>
<td>90%</td>
<td>3,603</td>
<td>72.3</td>
</tr>
<tr>
<td>100%</td>
<td>3,552</td>
<td>71.3</td>
</tr>
</tbody>
</table>

The gain from circumcision varies according different parameters. Illustrating Proposition 1 derived from the theoretical model in section 4, Figures 6 (a) and (b) show that the benefit from investing in circumcision is increasing in the cost of ART and decreasing in the cost of circumcision. Similarly, Figures 6 (c) shows that the benefit from investing in circumcision is decreasing in the marginal interest rate, as derived in Proposition 2. The latter indeed represents the opportunity cost of investments in prevention.

From Figure 6 (d), we also conclude that the return to investment in circumcision in Malawi is an inverted U-shaped function of the CD4 count threshold defining the moral duty to rescue. As explained in section 4, two opposite effects are at play. On the one hand, a high CD4 count threshold reduces the average infectivity of PLHIV, thereby reducing the need for prevention (Proposition 4). On the other hand, a high CD4 count threshold increases the expected cost of people not yet on treatment, thereby increasing the value of prevention (Proposition 5). In the case of circumcision in Malawi, the first mechanism seems to be stronger when the CD4 count threshold is high, and the second mechanism is stronger when the CD4 count threshold is low.

While our results show that prevention can be used to minimize the liability induced by the duty of rescue, they also show that the amount that will have to be paid remains very high. One reason for this is that a large part of the cost of the duty to rescue comes from people who are already infected. This ranges between 50 and 75 per cent depending on the country. Evidently, prevention can have no effect on the cost of treating those already infected. For poor countries with high-prevalence the total liability created by the duty to rescue is simply unaffordable, even after optimizing prevention to reduce the liability. In this case, the liability will have to be shared with the international community. The next section explains how this could be done without inadvertently introducing incentives that would distort behaviour.
Figure 6: reduction in the total cost of the moral duty of rescue from circumcision - 350 cells/mm$^3$ threshold.

6. Global Solidarity: Apportioning the Moral Duty to Rescue and Expenditure on HIV Prevention

Funding for HIV/AIDS from both domestic and international sources has grown rapidly over the past decade. By 2012 an estimated $18.9bn was spent on HIV programs in low- and middle-income countries. This was almost equally split between domestic and international sources, with the former accounting for 53 per cent.\textsuperscript{26} However, this global picture masks considerable variation between countries, depending on the size of their economy and the burden of HIV. A study by the Results

\begin{footnote}
\textsuperscript{26} The range of estimates is $16.6$–$21.2bn.
\end{footnote}
for Development Institute for PEPFAR reports that for nine high-prevalence African LMICs and LICs, if domestic resources rose at the rate of expected economic growth during 2012-16, dependence on external resources would vary within the range 84-99 per cent, (R4D, 2013). Thus, dependence on external sources of finance will be a feature of HIV/AIDS financing for many poorer high prevalence countries for a very considerable time. We therefore investigate how obligations might be shared between the citizens of poor countries and those of richer countries, and how this is best divided between the direct duty to rescue by means of treatment, and the optimal reduction in liability, by means of prevention.

Regardless of the share of external sources of finance in the total, two important principles should guide the apportionment of treatment relative to prevention as between donors and the affected countries in order to maintain incentive compatibility.

The first principle is that a broadly similar distribution of burden-sharing should be applied to treatment and prevention. If the international community accepts to finance a share $\sigma$ of the liability created by the treatment of future infection, it should also support at least a share $\sigma$ of prevention efforts. Only this strategy for dividing responsibility can avoid moral hazard. If, for example, donors entirely funded the treatment of future infections while African governments were to fund prevention, there would be too little incentive for the latter to invest in prevention.

The second principle is that both parties need to subject themselves to a commitment technology to avoid time-inconsistency. For example, were either party to accept responsibility for whatever residual funding was necessary, the other party would have an incentive to under-finance. Given the moral imperative of the duty of rescue, a commitment technology is particularly important, since otherwise each party will be tempted to leave the other with final responsibility. A simple and therefore easily enforceable commitment technology is continuous time pooled funding with specified matches to an agreed schedule.

Turning from these basic principles of incentive compatibility to how the overall burden should be shared, an adequate answer requires taking a normative position, and this is beyond the scope of the paper. We adopt a simple extension of the Bolam test, namely to infer the approximate division of responsibilities for the quasi-debt generated by the duty of rescue, from observed donor criteria in respect of past debt burdens and of the allocation of aid. In other words, we infer the implicit norms revealed by the community of donor practice, and accept that they are thereby to be judged reasonable.

The debt relief provided to poor countries during Jubilee 2000 provides one basis for determining whether the quasi-debt burden of HIV is too large to bear since the international community has already revealed its preference for assuming the burden itself. Several criteria have been used to determine the access and extend of debt relief. The exports criterion is arguably not relevant, since unlike international debt, some of the costs of delivering ART are in local currency. The debt-to-fiscal revenue target of 250 per cent is more pertinent, but perhaps the most straightforward is the rule-of-thumb proposed by the IMF whereby the debt of developing countries should not
breach 40 per cent of their GDP (Geithner, 2002; Belhocine and Dell’Erba, 2013). The analysis of table 2 shows that the moral duty to rescue imposes a financial liability which may be considerably higher than these thresholds, thereby warranting support from the international community.

An alternative criterion that might be used to determine the share that countries can reasonably assume derives from the Abuja target. In April 2001, African Union countries pledged to increase government funding for health to at least 15 per cent of their annual budget. The share of HIV in the overall burden of disease in a country could then be used to determine the share of its government health budget that should be allocated to HIV. This approach is likely to be flawed. Aspirational pan-regional target expenditure shares for particular sectors as liable to be biased upwards and so cumulatively incompatible. Only the actual national budget process truly reconciles competing claims. Further, applied literally, this method would provide the perverse incentive of discouraging tax collection. It is therefore necessary to adopt some ‘reasonable’ norm for tax collection rather than actual revenues. In columns (3) to (4) of table 4, we illustrate this approach by assuming that the tax norm is 15 per cent of GDP, since this was the norm used by the IMF in its HIPC calculations. With this assumption, the expected contribution of countries to the quasi-liability ranges from 16.8 per cent for Malawi to the entire 100 per cent for South Africa and Nigeria.

<table>
<thead>
<tr>
<th>Country</th>
<th>Cost in 2015 (% GDP)</th>
<th>Aggregated cost with incidence (% GDP)</th>
<th>Share of HIV in the total burden of disease (WHO)</th>
<th>Share that the country can afford</th>
<th>World Bank Classification</th>
<th>HIV Burden classification</th>
<th>Counterpart financing Global Fund</th>
</tr>
</thead>
<tbody>
<tr>
<td>Botswana</td>
<td>1.04</td>
<td>21.4</td>
<td>31</td>
<td>0.70</td>
<td>UMIC</td>
<td>Extreme</td>
<td>60</td>
</tr>
<tr>
<td>Kenya</td>
<td>0.34</td>
<td>9.1</td>
<td>14</td>
<td>0.31</td>
<td>LIC</td>
<td>Severe</td>
<td>5</td>
</tr>
<tr>
<td>Lesotho</td>
<td>1.95</td>
<td>73.6</td>
<td>40</td>
<td>0.91</td>
<td>Lower LMIC</td>
<td>Extreme</td>
<td>20</td>
</tr>
<tr>
<td>Malawi</td>
<td>3.33</td>
<td>80.3</td>
<td>25</td>
<td>0.56</td>
<td>LIC</td>
<td>Severe</td>
<td>5</td>
</tr>
<tr>
<td>Nigeria</td>
<td>0.15</td>
<td>6.8</td>
<td>8</td>
<td>0.19</td>
<td>Lower LMIC</td>
<td>Severe</td>
<td>20</td>
</tr>
<tr>
<td>South Africa</td>
<td>0.63</td>
<td>21.1</td>
<td>36</td>
<td>0.81</td>
<td>UMIC</td>
<td>Extreme</td>
<td>60</td>
</tr>
<tr>
<td>Uganda</td>
<td>0.72</td>
<td>21.3</td>
<td>15</td>
<td>0.34</td>
<td>LIC</td>
<td>Severe</td>
<td>5</td>
</tr>
<tr>
<td>Zimbabwe</td>
<td>1.81</td>
<td>38.7</td>
<td>26</td>
<td>0.58</td>
<td>LIC</td>
<td>Extreme</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 4 – Existing Methods to Determine the Share of the Quasi-Liability Supported by Countries - 350 cells/mm³ threshold

Development agencies have conventionally used income criteria for aid allocation and these could readily be used to share the responsibility of rescue. For example, those societies that meet the criteria for Low-Income Countries may be presumed to be too poor reasonably to bear more than a very small fraction of the costs of ART and HIV/AIDS in general. Such criteria could be combined with a measure of the HIV burden to determine the apportionment of the duty to rescue between high-prevalence poor countries and the international community. The eligibility rules of the Global Fund to Fight AIDS, Tuberculosis and Malaria are already based on such a combination. The counterpart financing required by the Global Fund is increasing with categories of revenue, from 5 per cent for low-income countries to 60 per cent for upper-middle income countries. Further,
upper-middle income countries only have access to funding if they face an extreme, severe or high disease burden. Similar rules could be applied to share the liability resulting from the duty to rescue. This method is illustrated in columns (6) to (8) of table 4 for our sample of countries. It is more demanding for the international community than the method based on the Abuja target.

Finally, instead of basing income and HIV burden criteria on arbitrary levels as it is currently done by the Global Fund, the sharing rule could be determined according to the actual behaviour of the international community. The revealed preference of donor countries can be captured by regression analysis. A simple regression of international spending for HIV per capita (in log) on GDP per capita of recipients country (in log) and on HIV prevalence shows that the international community is more inclined to finance HIV programs in poor and high prevalence countries (results and robustness checks are presented in Appendix B). Similarly, a regression of domestic spending for HIV per capita (in log) on GDP per capita (in log) shows that recipient countries are able to double their contribution to HIV programs if their income per capita doubles. The results of these regressions can be combined to determine the share that recipient countries can be expected to pay, both in term of financial capacity and in terms of the revealed preference of the international community (figure 7). According to this method, high-prevalence and poor countries like Lesotho, Zimbabwe, Malawi and Uganda would contribute to less than 20 per cent of financing needs. Countries that are richer or have a lower HIV prevalence would have to contribute more, up to 60 per cent for Nigeria for example.

![Figure 7 - Local funding percentage based on revealed preference of the International Community](image-url)
An important factor to be incorporated into the analysis is that donors and recipients face very different costs of borrowing. This affects decisions because, as shown in section 2, the net present value of the quasi-liability increases sharply with a lower discount rate. As a corollary, optimal prevention efforts are also increasing in the discount rate (Proposition 2 from the model). Hence, the share of the quasi-liability which should be paid by donors will be much higher if the marginal interest rate considered is that facing the donors (conventionally set at 3 per cent) rather than the one of the recipient (around 7 per cent). The ensuing optimal prevention level will also be higher. For something as fundamental to a society as the value it places on the future, we presume that the preferences of the government should take precedence over those of donors. We therefore assume that the marginal interest rate of recipient countries is the one that should be considered for determining the share to be paid by both parties. However, once this share is determined, the apportionment of each party’s contribution between prevention and treatment must also be determined and here there is no compelling reason not to allow the discount rate of each party to affect its choice. Hence, the net present value of future liabilities will be higher for the donor countries because their discount rate is lower. While the donor share of the total burden is determined by using the discount rate of recipient countries, a sensible way for donors to reduce the net present value of their contribution is to bias their spending towards early prevention.

We illustrate how this mechanism would work within our two-period model. We consider two countries: one recipient high-prevalence country labelled R and one donor labelled D. In order to simplify the maths, we leave prevention out of the analysis. In this case, the net present value of the duty to rescue for the recipient country is equal to:

\[ NPV^R = (1 - a_0)I_0 \pi^t + \frac{[(1-a_0)I_0 + (1-a_1)(a_0I_0 + N_0)] \pi^t}{(1+r^R)} \] (5)

Given its future flows of revenue, the recipient can only afford to pay an amount \( X \). The rest, \( NPV^R - X \), should be paid by the donor. The donor can determine the amount \( \Delta_0 \) paid in period 0 and the amount \( \Delta_1 \) paid in period 1, as long as the present value of its total contribution over the two periods is given by \( NPV^R - X \). Formally, its budget constraint is:

\[ \Delta_0 + \frac{\Delta_1}{(1+r^R)} = NPV^R - X \] (6)

Taking this constraint as given, it would be sensible for the donor to minimize the net present value of its own contribution:

\[ \min_{\Delta_0, \Delta_1} \Delta_0 + \frac{\Delta_1}{(1+r^D)} \] (7)

If we introduce the budget constraint given in equation (6) into this minimization function, we obtain:

\[ \min_{\Delta_1} NPV^R - X = \frac{\Delta_1}{(1+r^R)} + \frac{\Delta_1}{(1+r^D)} \]
As \( r^R > r^D \), the first-order condition of this minimization program is always positive implying that it is optimal for the donor to set \( \Delta t \) as low as possible. In other words, it should concentrate its budget on the financing of ART in period 0.

This effect is likely to be reinforced by the two-way relationship between the apportionment rule and the revenue of the recipient country. The share of the quasi-liability that the high-prevalence country is able to afford on its own will depend on its current and future levels of revenue. In turn, these revenues are likely to depend on how the cost of the duty to rescue is shared with the international community. If the international community bears most of the current cost and only future costs are left for the recipient, the latter could be able to invest in development (or borrow less), thereby increasing its future revenue and reducing the share of the liability that should be supported by the international community. This mechanism is strengthened by the high marginal interest rate in developing countries. Donors should therefore bear most of the current burden while recipients focus on development in the short-run and only pay their share of the burden in the long-run. A further benefit of this is that it provides an exit mechanism for donors over the long term. However, this qualification to the principle that the donor share of prevention should be the same as that of treatment is critically dependent upon a credible commitment technology. Depending upon the wider political context, no such technology may be possible.

The model can also be used to illustrate this moral hazard problem. Let us assume that the sharing rule is such that the donor country pays a share \( \sigma^{t1} \) of the cost of ART in period 1. In this case, the minimization programme that the recipient country uses to determine its optimal level of prevention will be affected. Because the cost of ART is shared in period 1, the first term in equation (2) is multiplied by \( (1 - \sigma^{t1}) \):

\[
\min_{\pi_0} (1 - \sigma^{t1}) [(1 - a_0) I_0 + (1 - a_1) (a_0 I_0 + N_0)] \pi_t + (1 + r^R) c(p_0)
\]

The resulting optimal prevention level is reduced compared to the benchmark model:

\[
p_0 = c^{t-1} \left( \frac{(1-\sigma^{t1})(1-a_1)}{(1+r^R)} \right) \frac{l_0(1-l_0)\pi^t}{l_0} < p_0^*
\]

One way to solve this moral hazard problem is for the donor also to pay a share \( \sigma^P \) of the prevention efforts in period 0. In this case, the second term in equation (8) is also multiplied by \( (1 - \sigma^P) \). With this adjustment, the optimal level of prevention is given by:

\[
p_0 = c^{t-1} \left( \frac{(1-\sigma^{t1})(1-a_1)}{(1-\sigma^P)(1+r^R)} \right) \frac{l_0(1-l_0)\pi^t}{l_0} = p_0^*
\]

If \( \sigma^{t1} = \sigma^P \), prevention efforts will be equal to \( p_0^* \), which proves the second principle.

In practice, this simple rule may be slightly affected by the fact that discount rates are different in donors and recipient countries. Because their discount rate is lower, the incentive for donors to
invest in prevention is higher (proposition 2 above). The donor country would therefore wish to increase its contribution to prevention beyond its share of treatment. However, in reaction, the recipient country would reduce its own contribution to prevention. Despite this, if the donor has to pay a large share of the cost of future infections, the benefit to the donor of the net increase in prevention may exceed the cost induced by the recipient’s underinvestment in prevention. In this case, the donor country may invest more in prevention than its share of treatment.

The model can be used to formalize this mechanism. If the donor country changes its contribution \( \sigma^P \) to prevention efforts, the reaction of the recipient country is given by equation (10). Given this reaction function, it is optimal for the donor to minimize the value of its contribution, taking its own discount rate into account.

\[
\min_{\sigma^P} \sigma^T[(1 - a_0)l_0 + (1 - a_0)(a_0l_0 + N_0)] \pi^T + (1 + r^D) \sigma^P c(p_0)
\]

In order to obtain a closed form solution to this minimization programme, we need to specify the cost function, \( c(p_0) \). Let us assume that it is a quadratic function of investment in prevention: \( c(p_0) = \pi^P p_0^2 \). With this assumption, we obtain that the optimal investment in prevention for the donor country is given by:

\[
\sigma^P* = \frac{2(1+r^R)\sigma^T-(1+r^D)(1-\sigma^T)}{2(1+r^R)\sigma^T+(1+r^D)(1-\sigma^T)}
\]

We have that \( \sigma^P* > \sigma^T \) if:

\[
\sigma^T > \frac{1+r^R}{(1+r^R)+(r^R-r^D)}
\]

If the marginal interest rate in donor country is 3 per cent and the marginal interest rate in recipient country is 7 per cent, the right-hand side of this inequality is equal to 93 per cent. In words, this tells us that if in period 1 the donor country has to pay at least 93 per cent of the cost of the moral duty to rescue, it is then optimal for it to invest more than its fair share \( \sigma^T \) in prevention. Indeed, because the donor has a lower discount rate, it values prevention more than the recipient. However, if the donor increases its contribution \( \sigma^P \) to prevention, it will be rational for the recipient to reduce its own investment in prevention. Nevertheless, when inequality (13) is satisfied, the former effect is stronger than the latter and it is optimal for the donor country to overinvest in prevention.

7. Conclusion

In this paper we have explored the implications of a moral duty of rescue for the treatment and prevention of HIV. While health economics conventionally adopts an ethical framework of Utilitarianism, this is radically at odds with prevailing moral norms in both developed and developing societies. Hence, in democracies where public policy can be expected to reflect prevailing norms, the precepts derived from conventional economic analysis may have little purchase. In
contrast to Utilitarianism, a moral duty to rescue is, we suggest, very widely recognized in most societies. Hence, its implications can have real purchase on public policy. Further, the distinctive features of HIV make the moral duty to rescue particularly pertinent: the decision whether or not to provide ART is tantamount to deciding whether a clearly specified group of identifiable people should live or die. It should therefore be possible to get broad ethical consensus on the need to keep HIV+ people alive by means of treatment, even if there is no such consensus on less sharply defined possible policies. In effect, recognition of the duty to rescue people who are HIV+ is moral minimalism, and as such is highly convenient for policy making in a democracy. The policies that follow from it are not dependent upon the imposition of a moral standard that is not widely shared.

While ethically minimalist, the duty to rescue has powerful and non-obvious implications. The first is that because ART treatment is an irreversible commitment which must continue for many years, almost all the costs of the decision lie in the future and so it incurs a liability analogous to a debt. We show that for some African countries this quasi-debt is very substantial, and indeed may make them debt-distressed. The moral duty to rescue cannot be absolute: it is conditional upon financial feasibility. However, it is a universal value, so where the fellow-citizens of HIV+ people are unable to accept such a duty due to their own poverty, the citizens of richer countries become bound by it. A reasonable allocation of the burden between poor countries with insupportable burdens and a much richer international community can be inferred from established practices both in respect of debt relief and aid allocation.

The key step in the paper has been to show that because the moral duty to rescue generates a substantial quasi-debt, it also generates the criterion by which to determine the minimum efficient level of expenditure on prevention. Prevention policies should be pursued at least up to the point at which expenditure on them reduces the debt burden sufficiently to minimize the overall cost of accepting the duty to rescue.

Finally, we show that whatever division of the treatment burden between the governments of poor countries and the international community is determined, that same division should apply to the funding of prevention. Any other assignment of the financial costs of prevention will generate moral hazard and so potentially give rise to inefficient resource allocation.

In this paper we have illustrated our analysis with applications to a few countries and specific strategies for prevention. As data improve, especially on the efficacy of prevention strategies, it will become possible to apply this approach to guide vital decisions about resource allocation.27

References


27 Our approach may soon also become pertinent for the treatment and prevention of cancer in OECD societies, as innovations in high-cost drugs turn many cancers into chronic conditions.


DART Trial Team DTT. Fixed duration interruptions are inferior to continuous treatment in African adults starting therapy with CD4 cell counts < 200 cells/microl. AIDS. 2008;22(2):237-247.


Knight, R., Small, W., Pakula, B., Thomson, K., & Shoveller, J. (2014). A scoping study to identify opportunities to advance the ethical implementation and scale-up of HIV treatment as prevention: priorities for empirical research. BMC Medical Ethics, 15, 54.


Schouten, Judith, Ferdinand W. Wit, Ineke G. Stolte, Neeltje Kootstra, Marc van der Valk, Suzanne G. Geerlings, Maria Prins, Peter Reiss, and on behalf of the AGEhIV Cohort Study Group. 2014. “Cross-Sectional Comparison of the Prevalence of Age-Associated Comorbidities and Their Risk Factors between HIV-Infected and Uninfected Individuals: The AGEhIV Cohort Study.” Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America, September.


from low-and middle-income countries. Bulletin of the World Health Organization, 92(7), 499-511AD.


Appendix A – methodology for estimates

UNAIDS 2014 files for SPECTRUM were asked on http://apps.unaids.org/spectrum/ and downloaded for Kenya, Lesotho, Malawi, Uganda, Zimbabwe, Botswana, Nigeria and South Africa.

After opening the UNAIDS 2014 for a given country: 12 steps are needed to produce the estimates for columns (4), (5), (8) and (9) of section 3:

1. The “Manager” is opened, the GOALS and RNM modules are selected and the timespan is extended to 2011-2050.
2. In GOALS, “configuration”, the target year is set at 2050.
3. In GOALS, “coverage”, prevention is set at 0 for all from 2015 for all interventions except circumcision (circumcision rates are in large part driven by other motives than HIV prevention and it is of course unrealistic to imagine that circumcised individuals would get uncircumcised).
4. In AIM , “program statistics”, PMTCT is set to 0 from 2015 onwards.
5. In AIM “program statistics”, adult ART is set at 81% from 2015 onwards.
6. In AIM “program statistics”, child ART is set at 81% from 2015 onwards.
7. In AIM, “eligibility for treatment”, no population are set to be eligible for treatment regardless of CD4 count (all boxes are unchecked).
8. In AIM, “eligibility for treatment” is set at 350 CD4/mm³ or 500CD4/mm³ depending on the scenario.
9. In RNM, “results”, the “number of people reached summary” is copy pasted in an Excel spreadsheet.
10. In GOALS “unit cost”, all unit costs are set equal to 0 except from the cost of first line ARV and second line ARV which are set equal to 1. This will allow us obtaining the number of adults on first and on second line therapy. The migration rate from first to second line therapy is set at 2.64% (Renaud-Théry et al., 2011).
11. In RNM “results”, “adult treatment costs”, the summary of costs is copy-pasted into spreadsheet. Given step 10, we can obtain the numbers on first line and second line ARV from this table. However, some precautions should be taken. First, it should be noted that the number on first line reported by RNM is in fact the total number of adults on ART (1st and 2nd line included). In order to obtain the real number of people on first line, one should take the number on first line reported by RNM minus the number on second line plus the number of children. The number of adults on second line is correctly reported by RNM.
12. The numbers of people on first and second line therapies are multiplied by the cost of ART in each country. These unit costs and their references are summarized in table A1. Columns (4) and (5) provide the result of this multiplication for the year 2015. For obtaining the numbers shown in columns (8) and (9), we first discount and sum the costs of first line and second line therapies between 2015 and 2050. Then, in order to take into account the costs which are likely to occur after 2050, we do a terminal value calculation. We divide the cost of the moral duty to rescue in 2050 by the discount rate to obtain the value of perpetuity. We
then discount this value of perpetuity and add the amount to the total discounted cost of the moral duty to rescue.

Table A1 – unit costs per intervention (Schwartländer et al., 2011)

<table>
<thead>
<tr>
<th>Services</th>
<th>Botswana</th>
<th>Kenya</th>
<th>Lesotho</th>
<th>Malawi</th>
<th>Nigeria</th>
<th>South Africa</th>
<th>Uganda</th>
<th>Zimbabwe</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UMIC</td>
<td>LIC</td>
<td>Lower LIC</td>
<td>LIC</td>
<td>Lower LIC</td>
<td>UMIC</td>
<td>LIC</td>
<td>LIC</td>
</tr>
<tr>
<td>ART - first line</td>
<td>704</td>
<td>343</td>
<td>430</td>
<td>343</td>
<td>430</td>
<td>704</td>
<td>343</td>
<td>343</td>
</tr>
<tr>
<td>ART - second line</td>
<td>2797</td>
<td>743</td>
<td>1157</td>
<td>743</td>
<td>1157</td>
<td>2797</td>
<td>743</td>
<td>743</td>
</tr>
<tr>
<td>Circumcision</td>
<td>64</td>
<td>64</td>
<td>64</td>
<td>64</td>
<td>64</td>
<td>64</td>
<td>64</td>
<td>64</td>
</tr>
</tbody>
</table>

Four more steps are needed to obtain the estimates without incidence which are displayed in columns (5) and (6) in table 2. In particular, the four channels of transmission modeled in SPECTRUM should be set equal to zero from 2015 onwards.

1. Sexual transmission is set to 0 by setting the “Number of partners” to 0 in GOALS, “behavior”.
2. Transmission for mother to child is set to 0 by setting abortion rate to 100 in AIM, “Program statistics”.
3. The force of infection for IDU is set to 0 in GOALS “Behavior”.
4. The number of new infections due to blood transfusion is set to 0 in GOALS “Epidemiology”

Then, the steps 11 and 12 are repeated to obtain estimates of the total cost of ART for this “no-incidence” scenario. It is worth noting that this procedure could not be undertaken for Kenya. For unknown reasons, setting sexual transmission to 0 from 2015 onwards for Kenya also change the incidence estimates before 2015, which is incoherent.

Appendix B – regression analysis

Preferences of donor countries are revealed by a simple regression of international spending for HIV per capita (in log) on GDP per capita of recipients country (in log) and on HIV prevalence. In columns (1) to (5), the sample is constituted of low- and middle- income countries for which information is available. In columns (6) to (10), the sample is restricted to sub-Saharan African countries. Regressions (1) and (6) are simple OLS regressions. In columns (2) and (7), we present the results of quantile regressions. In columns (3) to (5) and (8) to (10), we follow Colombo et al. (2014) by matching countries with their neighbours to control for unobserved characteristics that are similar in neighbouring countries. In columns (3) and (8), countries are matched with all their neighbours. In columns (4) to (9), countries are matched in pairs of neighbours. In columns (5) and (10), each country is randomly matched with one of its neighbours; this process is then repeated for 100
random matching and coefficients are averaged. In each matching method, standard-errors are clustered at multiple levels to account for the fact that countries may have multiple neighbours and may be neighbours of different countries. Results are robust to the different estimation methods.

Table B1: revealed preferences of donor countries (standard-errors in parenthesis)

<table>
<thead>
<tr>
<th>Matching methods</th>
<th>Quantile regression</th>
<th>Neigh.</th>
<th>Pair matching</th>
<th>Random matching</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLS (1)</td>
<td>Log(GDP per cap.)</td>
<td>-0.622***</td>
<td>-0.493***</td>
<td>-0.552***</td>
</tr>
<tr>
<td></td>
<td>(0.153)</td>
<td>(0.166)</td>
<td>(0.163)</td>
<td>(0.194)</td>
</tr>
</tbody>
</table>

|                  | HIV prevalence    | 20.83*** | 16.25*** | 20.29*** | 22.06*** | 21.72*** | 16.35*** | 15.12*** | 15.68*** | 17.24*** | 16.86*** |

|                  | Constant          | 4.229*** | 3.715*** | 2.446**  | 2.640*   | 2.980*** | 2.380*   | 1.958*   | 0.0554   | 1.660*   | 1.786**  |
|                  | (1.098)           | (1.275) | (1.237)  | (1.468)   | (1.019)  | (1.189)  | (1.067)  | (1.073)  | (0.965)  | (0.634)  |

| Observations     | 93                | 93      | 357       | 550       | 164      | 42       | 42       | 190      | 304      | 76       |
| Countries        | All               | All     | All       | All       | All SSA  | SSA      | SSA      | SSA SSA  | SSA SSA  | SSA SSA  |

Table B1: revealed preferences of donor countries (standard-errors in parenthesis)

The financial capacity of recipient countries is measured by regressing their domestic spending for HIV per capita (in log) on their GDP per capita (in log). The structure of table B2 is similar to table B1. Results are robust to the different estimation methods. Figure 7 is based on columns (6) of tables B1 and B2.

Table B2: financial capacity of recipient countries (standard-errors in parenthesis)

<table>
<thead>
<tr>
<th>Matching methods</th>
<th>Quantile regression</th>
<th>Neigh.</th>
<th>Pair matching</th>
<th>Random matching</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLS (1)</td>
<td>Log(GDP per cap.)</td>
<td>0.923***</td>
<td>0.822***</td>
<td>1.154***</td>
</tr>
<tr>
<td></td>
<td>(0.134)</td>
<td>(0.144)</td>
<td>(0.202)</td>
<td>(0.223)</td>
</tr>
</tbody>
</table>

|                  | (1.053)           | (1.126) | (1.841)   | (1.89)   | (1.719)  | (1.475)  | (2.05)   | (1.351)  | (1.753)  | (1.549)  |

| Observations     | 118               | 118     | 466       | 734       | 201      | 45       | 45       | 201      | 324      | 78       |
| Countries        | All               | All     | All       | All       | All SSA  | SSA      | SSA      | SSA SSA  | SSA SSA  | SSA SSA  |

Table B2: financial capacity of recipient countries (standard-errors in parenthesis)