HIV Testing: A Trojan Horse?

Stéphane Mechoulan, University of Toronto
HIV Testing: a Trojan Horse?

Stéphane Mechoulan*

*University of Toronto, s.mechoulan@utoronto.ca

Copyright ©2004 by the authors. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher, bepress, which has been given certain exclusive rights by the author. Topics in Economic Analysis & Policy is produced by The Berkeley Electronic Press (bepress). http://www.bepress.com/bejeap
HIV Testing: a Trojan Horse?

Stéphane Mechoulan

Abstract

The consequences of HIV testing are unclear. Some infected individuals, assuming they behave selfishly, would tend to increase their number of partners. Meanwhile, non-infected ones or those ignorant of their status would decrease theirs, the result of which, on the equilibrium level of infection, is uncertain. Simulations from a simple dynamic model show how to generate the Philipson-Posner conjecture, i.e., that disclosure of HIV status may result in higher disease prevalence. In this benchmark case, testing would also lower welfare. Those results, however, appear to be fragile. In particular, very little altruism seems needed for testing to become beneficial, and the public health literature tells us that a large proportion of individuals behave altruistically when tested positive and appropriately counseled. Beyond the mere availability of testing, the findings further suggest combining existing prevention measures with universal or mandatory testing to help eradicate the disease.

KEYWORDS: HIV, testing

*I would like to thank Gadi Barlevy, Ted Bergstrom, Michael Denny, Jeff Ely, Pierre-Yves Geoffard, Li Hao, Ig Horstmann, Michael Smart, Chris Taber, Burton Weisbrod, Ralph Winter, as well as Dino Gerardi and Nicolas Sahuguet, seminar participants at Rotterdam’s second World Conference in Health Economics, and particularly David Bishai, Stefano Bertozzi, and Mead Over for useful comments and discussions. Tomas Philipson made several suggestions that considerably improved an earlier draft. Finally, the paper has benefited from the numerous and helpful insights of two anonymous referees. Financial support from the Northwestern University Weinberg College of Arts and Sciences and from the Northwestern University Graduate School is gratefully acknowledged. All errors and omissions are mine. Affiliation: Department of Economics, University of Toronto, 150 St. George Street, Toronto, Ontario M5S 3G7, Canada. E-mail: s.mechoulan@utoronto.ca
1 INTRODUCTION

Since its appearance in the early 1980’s, the HIV virus has caused worldwide devastation. At least 40 million people have been infected, around 95% of them living in developing countries. It is believed that over twenty million have died of AIDS so far.¹ Short of an eradication method, public policy is crucial in fighting the epidemic and many steps have been taken, like subsidization of condoms and sex education.

The availability of information about one’s own serological status, potentially brings a new tool to contain the plague. Specifically, the key questions regarding testing are whether it contains the spread of HIV and whether it improves welfare. There are several reasons why the answers to those questions are ambiguous. First, we generally do not know if people behave in an altruistic or in a selfish way.² Second, the adjustment of susceptibles, i.e., not infected, or that of people ignorant of their status to the new situation that testing brings is of unknown magnitude. In other words, disclosing infection status may cause a selfish increase in risky behavior by some of the infected while it causes more conservative behavior by the uninfected.

In the economics literature, Philipson and Posner [1993] offer an exploratory discussion on HIV testing which casts doubts upon its benefits. Boozer and Philipson [2000] estimate the impact of HIV testing on subsequent sexual behavior, yet they cannot assess the ultimate effects on disease prevalence (that is, the fraction of the population infected by HIV) and welfare with their data.

The present paper constructs a dynamic theoretical model to consider how testing combined with individual choices on sexual exposures will affect HIV prevalence and welfare under both selfish and altruistic preferences. The structure chosen is a simple, stylized two-period framework where the key parameter is the number of risky sexual exposures.³ It builds on previous work by Kremer [1994

¹ UNAIDS report [2003].

² Here, it may be necessary to warn against a common pitfall. Altruism, for a HIV-positive person, does not mean the latter will stop unprotected (or risky) exposures altogether. An infected individual who engages in some unprotected exposures but still fewer than what he would have if there were no disease displays a behavior that incorporates an altruistic component.

³ Allowing for people to test frequently is hard to model in a dynamic setting. That is why I consider a situation where everyone tests once, in the middle of their two-period life, and compare it to the case where no testing is available. I conjecture that any qualitative result that arises from one-shot testing should go in the same direction as we move toward perfect information for every individual at each instant.
and 1996] and, more generally, the paper draws from a vast body of empirical public health literature on testing issues.

The base-case model is meant to generate and provide more insights into the Philipson-Posner conjecture, i.e., that the availability of HIV testing may increase, rather than reduce the spread of AIDS. Individuals are initially assumed to behave selfishly. Hence, in the case of disclosure, infected individuals increase their activity compared to what they would have done had the test not been implemented. Those who have tested negative or who have not tested yet would indeed compensate for this effect by reducing their desired number of risky exposures. However, for them, the adjustment has a cost in terms of foregone utility. The simulations of this paper show that those individuals would indeed correct their number of partners downward; but, in this base-case, such a modification is never enough for the prevalence not to grow. Similarly, welfare decreases because prevalence is higher in the case of disclosure, while expected lifetime sexual activity is also lower. This can be seen as a variation of the classical Hirshleifer effect.

However, this benchmark case is extreme and the paper investigates the robustness of the Philipson-Posner conjecture in several ways. The key point is that only a small proportion of altruists is needed to offset the perverse effect previously mentioned. I also consider how many such individuals would be necessary to eradicate the disease within a universal testing program. Second, as the prevalence increases, screening becomes a more important tool of protection; consequently, the assumption of random mixing becomes less credible. People who can exhibit a negative test result would be more likely to match with each other. Interestingly, welfare results change in an ambiguous way when susceptibles can disclose their negative status. Third, infected people tend to live a shorter time than susceptibles, and small differences in life duration prove to have dramatic effects. Therefore, it turns out that the Philipson-Posner conjecture is fragile in the sense that it is sensitive to minor changes in the setup that may produce it. Overall, it is more than likely that testing programs are highly efficacious.

The remainder of the paper is organized as follows. Section 2 briefly reviews the relevant literature and provides a motivation for the research. Section 3 describes the base-case model that generates the Philipson-Posner conjecture. Section 4 presents the simulation results of the model and their interpretation.

---

4 So, mechanically, for those two reasons, the odds that one random partner available for a risky exposure is infected – what I refer to as “practical prevalence” – becomes higher. The distinction between this “practical prevalence” and the general prevalence in the population is indeed critical.

5 That is, the impossibility of trading risks that have been resolved.
Section 5 discusses the robustness of the findings and shows how that conjecture may be overturned. Section 6 concludes and offers directions for future extensions.

2 REVIEW OF THE LITERATURE AND MOTIVATION

2.1 THE ECONOMIC APPROACH TO THE HIV EPIDEMIC

There is strong support for the economics approach to epidemiology which assumes behavioral adjustments to a changing environment. Therefore, I shall not review the traditional epidemiology methodology, which does not take incentives into account.  

Models of HIV spread can take different views about integrating behavioral choices; Gersovitz and Hammer [2003] provide a thorough discussion of the different elements at stake. Concisely, the fact that people have at their disposal either protection from risky behavior, e.g., condoms or screening – or a reduction in sexual activity itself shows the kind of arbitrage economics can model when studying so-called rational epidemics. Such a range of options also gives a flavor of the necessary simplifications that any analysis will have to make.

The starting point of the approach is to make the individual’s decision dependent on the prevalence of the disease. Specifically, the endogenization adopted by Kremer [1996] can be considered the block I build upon.

2.2 THE THEORY OF HIV TESTING

According to Philipson and Posner:

“Policy makers and public health scholars alike believe that widespread testing for HIV is essential to reducing the incidence of AIDS (...) availability of HIV testing may increase, rather than, as generally believed, reduce the spread of AIDS.”

---


7 This is reminiscent of intensive/extensive margin problems in the labor literature.

8 See Geoffard and Philipson [1995]. The critiques on the ground of rationality being problematic to such situations are reviewed in Philipson and Posner [1993].

9 Philipson and Posner [1993, chap. 3, p. 84].
The reason is, to summarize their arguments, that enough individuals who test positive will behave selfishly and thus have incentives to increase the number of their partners and, if possible, talk their partners out of adopting protective measures, to counterbalance the opposite effect, from altruistic infected people; thus there is no longer a clean theoretical prediction of the net effect of voluntary testing on the spread of AIDS. The rationale for an increase in sexual activity by selfish infected individuals is the notion that they are in an absorbing state and therefore have “nothing to lose.”

Philipson and Posner’s informal prediction is that prevalence is likely to increase. However, they emphasize that their analysis is made in a static framework. One of the purposes of this paper can therefore be viewed as an attempt to give a first answer to the following puzzle and the conjecture that the authors make about it:

“The underlying problem is that the dynamics of the process are highly complex. It is relatively clear that the availability of testing enables some people to obtain risky sex that they could not obtain were testing unavailable, but it is uncertain what happens next, when the market rearranges itself in response to the outcome of the test results. Hence, were it proposed to forbid HIV testing in order to reduce the spread of the disease, we would emphasize the lack of firm proof that such testing has in fact increased that spread. But (...) we are entitled to emphasize (...) the reasons for believing that testing is unlikely to reduce, and may increase, the spread of the disease.”

At an intuitive level, the reason why the answer is generally ambiguous depends on the distribution of parameters (ex: proportion of sexually active, promiscuous individuals), technology (ex: transmission rate per contact), and tastes (ex: proportion of selfish agents). The empirical literature gives some idea of the latter.

2.3. Empirical Results of Testing

From the epidemiology and public health literature, the conclusion that emerges is that on average, individuals who have tested positive seem to decrease their number of partners or adopt protection following testing. Yet, several caveats

---

10 ibid., p. 106. See also Coates et al. [1988].

11 ibid., p. 108.

12 See Weinhardt et al. [1999], Grinstead et al. [2001], Grésenguet et al. [2002], for recent studies.
should be emphasized to understand why we should not take these results at face value.

The first one is that those being tested generally do so voluntarily, and are always assumed to be more altruistic than the rest of the population [Cleary, 1991]. Second, in nearly all these studies, there is an implicit yet strong assumption of exogenous attrition – and we may suspect those people who drop out of the study to be less altruistic than the rest. The third caveat is that of reliability. As a corollary, finally, these studies almost always have testing accompanied by counseling: for obvious ethical considerations, it is thought unacceptable to have a controlled group of people tested but not counseled; for the researcher, it is thus hard to disentangle the two effects.

In the economics literature, Boozer and Philipson [2000] analyze a Californian controlled experiment where some people were tested and others were not, and the variable of interest was the behavioral impact of the test result. Their conclusion is that testing programs appear to do little in terms of changing high-risk group behavior regarding number of partners and protection. Why individuals in general would behave in this way may be caused by “rational fatalism.”

In the following, I first present a model where all individuals behave selfishly; this is one of the assumptions I relax later on to test the robustness of the base-case results.

**3. THE BASE-CASE MODEL**

**3.1 ASSUMPTIONS**

The following discrete framework is intended to establish how the Philipson-Posner conjecture may be generated: it should be considered a point of departure.

A cohort is born every period and is always replaced by another of the same size; I assume for now that people live for two periods, i.e., “young” and “old,” and that age is only privately observed. ¹⁵

---

¹³ “Social desirability bias may be higher among seropositive individuals who may feel under greater pressure than seronegative individuals to report behavior change after intensive counseling about risk reduction.” [De Zoysa et al., 1995, p. 598].

¹⁴ Gersovitz and Hammer [2003] provide a clear discussion of this phenomenon.

¹⁵ While one may note the resemblance of this setup with that of an overlapping generation (OLG) model, it is worth emphasizing that the two groups in any single period (the young and old) need not be thought of as generations. And in particular, the framework presented here has nothing to do with the transmission of HIV from mother to child, a major means of transmission in developing countries.
In each period, risk-neutral individuals may engage in a risky behavior by choosing a number of unprotected sexual contacts (or exposures). This model will only consider the number of exposures to the disease as the variable of interest. Formally, the level of exposures is the only endogenous variable in the agents’ utility function $u(.)$, assumed to be differentiable. This function could be different for old and young, but for simplicity it is the same.

I disregard protection issues, yet it is possible to view protected (costly) behavior and risky behavior as different goods, implying different, separate markets. Hence, a reduction in the number of risky exposures would correspond to a substitution toward risk-free behavior rather than a reduction in the sexual activity itself. Regarding the spread of the disease, the two interpretations are observationally equivalent if protection is fully efficient.

The assumptions regarding the disease dynamics are as follows. Contact among two individuals, one infected, one susceptible, is the only way for the disease to propagate. One does not get sicker by exposing oneself several times to the virus and the disease affects everybody the same, with no recovery (i.e., absorbing state). Whether one is infected or not does not affect the utility derived from exposures.

At the start of the second period, if testing is introduced, individuals get the information whether they carry the disease or not. The test conveys the correct information with certainty. The assumption that testing is mandatory, if the test is costless, does not affect the outcome: should testing be available in the population, everybody tests since it is in each individual’s best interest to do so, and this is common knowledge. For now I assume no possibility of test result disclosure, which will be relaxed later.

During the second period, individuals may engage in risky behavior again. When an individual carries the disease and knows it, in the benchmark case, he behaves “selfishly.” That is, infected individuals will choose the maximum number of exposures they would have chosen if there were no disease since I assume there is no risk for them anymore. To stress the distinction, the number of exposures chosen in this case will have a max subscript.$^{17}$ Then, when death takes place, disutility from the disease for those infected is in the form of a cost

---

$^{16}$ Otherwise, in this model, anybody has an incentive to match with the youngest partners. In equilibrium, everybody should meet a partner of the same age (classical assortative matching result), and the disease would be rapidly eradicated.

$^{17}$ Length of life is initially assumed independent of whether one carries the disease; I will later address the case of death at the end of the first period with an exogenous probability conditional on infection.
incurred at the time of death, call it \( c \). Note that \( u(. \) will be expressed in the same unit as \( c \).

At the time when matching takes place, individuals cannot see whether their partner is exposing for the first, second, ... or last time. But, I make the key simplifying assumption that an individual infected during one exposure cannot be infectious to others until the following set of rounds, that is, the next period. The rounds are sufficiently close in time so that the discounted value of the next period of life from a young person’s perspective is assumed to remain constant over the rounds. This discount factor is the same for everyone and called \( \delta \).

Prevalence, that is the proportion of individuals infected in the population as a whole, is called \( P \). Practical prevalence, that is the odds that someone picked at random from the pool of potential partners is infected, is called \( W \). Finally the probability that one susceptible gets infected after contact with an infected person, i.e., the transmission parameter, is called \( \beta \). The value of these parameters is common knowledge.

Most of the analysis presented in the next section characterizes equilibria where practical prevalence – endogenously derived from the young and old individuals' decisions – is constant over time. Occurrence of a steady-state in the case of such a disease as AIDS is one of the fundamental results of economics-based epidemiology.

Finally, a natural question can be raised about the meaning of the non-disclosure case when people maximize their lifetime utility function taking into account the current level of prevalence in the population. Obviously, to get this information, some testing must occur. The information could come from a representative sample of the population and the results not be disclosed to the tested individuals; or, more plausibly, they would be disclosed but the sample would be small enough to consider the population as a whole as non tested.

### 3.2 Matching Process

The matching process is defined in the following way. Let the population be divided into \( n \) groups, ranked by increasing number of desired exposures (a group is made of individuals who desire the same number of exposures). Let the maximum number of desired exposures be \( i_n \). This will be \( \text{Argmax } u(\cdot) \) if \( u \) has a maximum. Let \( \alpha_k \) be the percentage of people from group \( k \) in the population. A set of \( i_n \) rounds is organized in each period to ensure that those individuals who want \( i_n \) exposures will be “served.” So \( i_n \)-people will be the only ones present at every round of the period. During each round, people from all groups show up and mix randomly with each other. Let the number of exposures of people in group \( k \) be \( i_k \). On average, individuals from group \( k \) are present at any meeting with probability \( i_k / i_n \). The probability to meet an individual of group \( k \) is
Therefore, each individual will get his desired number of partners during the set of rounds, which is the purpose of this stylized setting.\textsuperscript{18}

Consequently, if prevalence in, say, group $k$ is $Y_k$, \textit{practical prevalence} $W$ (the odds of encountering an infected individual from a random draw among those who want an exposure) is:

$$W = \frac{\sum_{i=1}^{n} (i_1 \alpha_i) Y_i}{\sum_{i=1}^{n} (i_1 \alpha_i)}$$

(see Kremer [1994 and 1996]).

This completes the setup of the base-case model.

3.3 \textbf{Characterization of the Equilibria}

Let us call $i$ the number of partners when young and $j$ the number of partners when old.\textsuperscript{19} From the above, the model will allow for at least two and at most three groups: simply the young and the old if no disclosure of health status is provided and, if it is provided, the young, the infected old and the uninfected old.

\textbf{3.3.1 In the case where people receive the information about their status:}

Starting from the beginning of the second period of life, an uninfected old solves:

$$\text{Max}_{j} u(j) - [\text{Probability(getting infected with } j \text{ exposures})] \times c,$$

or:

$$\text{Max}_{j} u(j) - [1 - (1 - \beta W)^j] c.$$

An appropriate choice of $u$ will ensure that $j \geq 1$, which amounts to the following condition:

$$u'(j) = - \ln(1 - \beta W)(1 - \beta W)^j c.$$  

\textsuperscript{18} And when the size of groups is large the odd/even problem becomes negligible.

\textsuperscript{19} The number of exposures should be an integer, but for tractability I will treat it like a continuous variable.
Note that the probability that an old susceptible will be infected eventually is 

\[ 1 - (1 - \beta W)^j. \]

Let us call \( V \) the indirect utility function:

\[ V(j^*) = u(j^*) - [1 - (1 - \beta W)^j]c. \]

A knowingly infected old sets \( j = j_{\text{max}} \). \( V(j_{\text{max}}) = u(j_{\text{max}}) - c \). This is a constant.

The young solve

\[ \text{Max}_i u(i) + \delta[(1 - \beta W)^j V(j^*) + (1 - (1 - \beta W)^j) V(j_{\text{max}})]. \]

A steady-state equilibrium in this case is given by a triplet of numbers \((i^*, j^*, W^*)\) such that:

\[ u'(j^*) = -\ln(1 - \beta W) (1 - \beta W)^j i* c \]
\[ u'(i^*) = -\delta \ln(1 - \beta W^*) (1 - \beta W^*) i* [V(j^*) - V(j_{\text{max}})] \]
\[ W^* = (1 - (1 - \beta W)^j i_{\text{max}}) / [(1 - (1 - \beta W)^j i_{\text{max}}) + (1 - \beta W^*) i^* (j^*) + i^*] \]

### 3.3.2 In the case where people do not receive the information about their status:

I proceed similarly. The condition of the old becomes:

\[ u'(j) = -\ln(1 - \beta W) (1 - \beta W)^j i c. \]

This is because the old have to take into account the chances that they got infected from the exposures of their youth. Let \( j^* \) be the associated maximizer given \( i \) and \( V(j^*) \) the associated indirect utility function.

The young solve:

\[ \text{Max}_i u(i) + \delta V(j^*, i). \]

---

\(^{20}\) \( W, \beta, c \) enter the indirect utility function as parameters; in the following, I will just avoid putting them to soften notations.
This time:

\[ V(j^*, i) = u(j^*) - [1 - (1 - \beta W)^{i+j^*}]c. \]

A steady-state equilibrium in this case is given by a triplet of numbers \( (i^*, j^*, W^*) \) such that:

\[ u'(j^*) = -\ln(1 - \beta W^*)(1 - \beta W^*)^{i^*+j^*}c \]
\[ u'(i^*) = -\delta \partial V(j^*, i)/\partial i|_{i^*} = -\delta \ln(1 - \beta W^*)(1 - \beta W^*)^{i^*+j^*}c^{21} \]
\[ W^* = (1 - (1 - \beta W^*)^{i^*})^{j^*}/(j^* + i^*) \]

### 4 Solving the base-case model

Obviously, both systems are non-linear and cannot be solved analytically.\(^{22}\) Therefore, I now proceed with numerical simulations. At this point, it is useful to introduce a functional form for \( u(.) \). I focus on the case of linear quadratic functions, \( i.e.\), \( u(x) = \theta x - \psi x^2.\)\(^{23}\) One would like to think that exposures increase utility at first, and then decrease it beyond satiation. In particular, this convenient functional form makes the maximum number of possible partners transparent, namely \( j_{\max} (= i_{\max}) = \text{Argmax } u(x) = \theta/2\psi. \) Since a linear quadratic form implies a concave inverse U-shaped function, this maximum is unique. It is very convenient for testing the robustness of the analysis to change functional forms easily, and

---

\(^{21}\) This comes from the envelope theorem.

\(^{22}\) It turns out a model combining systems (2) and (3) by introducing an exogenous probability of being tested does not provide meaningful comparative statics. The reason for this is that such a model involves four equations (maximization problem of the young, of the tested negative old, of the non-tested old, and the prevalence defining equation). When applying the implicit function theorem to see the effect of an increase in the probability of testing on the activity of for example the young, there are too many cross partials. Although signing most of them is fairly easy, the outcome is ambiguous: economic reasoning does not tell which are first-order effects and which are not.

\(^{23}\) This is the same form as in Kremer [1996].
this can be done best by setting $\psi = 1/2$ and by making $\theta$ vary as in this case $\text{Argmax } u(x) = 0$. It has the very simple interpretation that it is the highest number of sexual encounters desired by the individual, in the absence of cost. I shall often refer choosing $\theta$ as the “max strategy” below.

For simplicity, I set the indirect utility of the infected $V(j_{\text{max}})$ to be zero. Finally I set $\beta = 0.033$ (drawing from Geoffard and Philipson, 1995) and $\delta = 0.8$.  

4.1 Description of the results

The range of values chosen for the parameter $\theta$ covers all illustrative cases: from low enough for the disease not to be sustainable, to high enough for testing to have virtually no effect. Can we think of $\theta$ as a realistic number of exposures? Because of the stylized nature of the model, $\theta$ should be best thought of as an abstract indicator of sexual activity in a population. Let us emphasize here that the model is meant to provide qualitative answers, and not to be used for a calibration exercise.

Although the following discussion focuses on steady-state equilibria, one word should be given on the transition dynamics. Because of the nature of a behavior-dependent epidemic, future infection levels depend on today’s expectations. Hence, we may run into the usual self-fulfilling prophecy and multiple equilibria problems. In this model, the myopic transition dynamics with $E_t[j_{t+1}] = j_t$ where $E[.]$ stands for the expectation operator implies a convergence to a non trivial steady-state, which is a desirable property for any model of economic epidemiology. The absolute speed of convergence here has no meaning because of the modeling assumption that infected individuals are not infectious until the following period (if then). We can just observe that the convergence is as rapid under both regimes.  

Let us now deal with the steady-states. For “low” values of $\theta$, it is found that the disease cannot sustain itself. For example, for $\theta = 50$, both information specifications yield the same equilibrium where $W$ is zero and everybody chooses $i = j = \theta$.

---

24 This is to keep the relative cost of the disease constant across $\theta$.

25 This corresponds to an implied discount rate equal to 25%. A higher $\delta$ would not affect the results qualitatively.

26 Illustrative examples are available from the author.
Things change when one plugs in “high” values for $\theta$. Figure I and II below show the different prevalences and number of partners for different values of $\theta$.

**Figure I**

**Figure II**
4.2 INTERPRETATION

First of all, the idea that the disease can propagate only when at least part of the population has a high number of partners\(^{27}\) is confirmed by the simulations. The intuition for this is that when an individual who is infected has a small number of partners, he is quasi lost from the point of view of the disease, since the odds that he will transmit it is close to zero\(^{28}\), so everybody can reach his bliss point (setting i or j equal to \(\theta\)).

The reason why the prevalence of the disease is zero in such cases is that for the disease to sustain itself, each infected individual must have on average more than \(2/\beta\) partners\(^{29}\) (one must take into account the possibility that two infected individuals meet, but this is negligible when prevalence is very low). This is because people who get infected during the second set of rounds are not infectious, and because there is a 0.5 probability that one random partner is old. Then, for the disease to grow, any infected must meet more than twice as many partners as the inverse of the transmission rate (which measures how many partners on average are necessary for the disease to propagate). In other words, on average, each infected person must transmit the disease to at least one young.

Another indication that the model yields consistent results is found by observing that the higher the maximum number of partners, the higher the prevalence.\(^{30}\) In this model, since the young who carry the disease are not infectious until the next period, and since people who get infected when old will not transmit the disease, the practical prevalence will never be higher than 0.5 by construction. So, it is normal that practical prevalence converges to 0.5 as \(\theta\) gets bigger. If we look at prevalence P, this means that eventually, nearly everybody is going to be infected. This happens for \(\theta\) in the range of 150 and up.

The number of partners and prevalence tend to be the same in both cases as \(\theta\) increases, which also makes sense. When people express a higher preference for sexual activity, they will tend to take more risks. This is a case in which testing has a very limited impact as it does not significantly alter people’s behavior.

Another interesting aspect of these simulations is that they shed light on the “multiplier effect” of average sexual activity in a population. Seemingly small, innocuous changes in \(\theta\) can produce dramatic changes in terms of prevalence, as

\[^{27}\) See Over and Piot [1993].

\[^{28}\) See Kremer [1994] for more on this point.

\[^{29}\) The minimum on the x-axis of the simulation graphs, 61, is evidently the critical value \(2/\beta\).

\[^{30}\) It is true for both total prevalence (P), and practical prevalence (W).
shown in figure II. This means that if we get to know this parameter in a population and it is found to be only a little higher than what is necessary to sustain the disease, the action targeted to reduce it will be very powerful, assuming it is somewhat elastic with respect to public intervention. Conversely, for higher ranges of $\theta$, the same action would have very limited success.

This multiplier effect is analogous to that induced by herd immunity in a population. If enough individuals in a community are immunized against a certain disease, then it is more difficult for that disease to get passed between those who are not immunized. Depending on the initial level of immunization, minor changes in the proportion of those immunized can have enormous effects on the overall caseload of infected.\footnote{I thank an anonymous referee for pointing out this interesting analogy.}

Whether within the testing or the non-testing case, the fact that the number of partners is higher among the young is a direct consequence of discounting. This effect should be and actually is found to be magnified in the testing case. In the latter case, it makes more sense for a non-infected individual at the start of the second period to cut down on exposures because there is a better chance that he will not be infected by the end of his life (since the “weight” of previous exposures when he was young disappears when computing the odds of being infected in the end).

What matters most is that the lifetime number of partners in the non-testing case is always higher than the expected lifetime number of partners in the testing case. This is because in the testing case, the pool of partners is worsened by the fact that all the knowingly infected old play the “max strategy.” It leads the young and susceptible old to being more “cautious.” The expected marginal increase in sexual activity conditional on testing positive, in the testing case, is second order. The simulation shows us individuals never fully compensate for the increased risks by reducing the number of exposures enough to make the probability of getting infected eventually equal to what it is in the non-disclosure case. This would be too costly to do in terms of exposures, and hence a higher $W$ (which is obvious) but also a higher prevalence $P$ in the testing case.

What about welfare implications? By plugging into the indirect utility functions the values obtained in the simulations, it is possible to compare welfare under both disclosure and non-disclosure for a same value of $\theta$. 

\footnote{I thank an anonymous referee for pointing out this interesting analogy.}
Figure III tell us that the non-disclosure equilibrium is better. This should be clear from the previous discussion as well. In the disclosure case, only those who test positive are better off *ex post*. However, this effect is never sufficient to increase *ex ante* welfare. Note that again, for very high values of $\theta$, disclosure does not matter much: welfare tends to be the same in both cases.

5 **ROBUSTNESS OF THE BENCHMARK RESULTS**

The previous analysis focuses on extreme assumptions that tend to bias the perverse effect upward. As it turns out, the above results are very sensitive to minor changes in the setup of the model and can be easily overturned. In the following, I show how disclosure of a test outcome and a higher mortality rate for infected individuals may contribute to mitigating those benchmark results. Of course, the most critical assumption is that of selfish behavior for the entire population. I investigate how much it should be relaxed to reverse the previous conclusions and how much altruism would be necessary to eradicate the disease. I analyze the three features separately.
5.1 Disclosure of Test Results, Monogamy, and Partial Assortative Serological Status

One should bear in mind that since infected partners are more likely to meet with other infected partners than the rest of the population, forecasts based on random matching will tend to bias prevalence estimates upward. Individuals presumably screen their partners better as a reaction to increased prevalence. To see how this could change the benchmark results, I introduce disclosure to partners. If the old susceptible can show their health status to their partners, there will be assortative matching for them in the sense that they will pair together and will withdraw from the rest of the population (young and infected old). Therefore, all the old set their number of partners to $j_{\max}$. The algebra is a straightforward adaptation of the first system of equations, and the results from the non-testing case are reproduced for comparison in Figure IV.

In the disclosure-to-partners case, it is obvious that practical prevalence in the general pool increases (recall that the old susceptibles can now match in a different pool). It can also be shown that the number of young who become infected also increases: despite the worsening of the pool (higher $W$) the decrease in $i$ is not sufficient to prevent an increase in the infection of the young. Yet, in terms of the proportion of eventually infected $P$, figure IV tells us the results depend on $\theta$. For small values of $\theta$, the negative effect dominates, but for high $\theta$, the certainty for the few non-infected old to escape infection more than compensates for the additional infection probability when young.

![Figure IV](http://www.bepress.com/bejeap/topics/vol4/iss1/art18)
What about welfare?

![Welfare Difference, Testing + possible disclosure to partners vs Non-Testing Equilibria](image)

Figure V

The results are again ambiguous, which is not surprising given the previous discussion. For low $\theta$, welfare is higher in the non-testing case and conversely for higher values of $\theta$. The results are also more sensitive to the value of $\delta$: the more discounted the future, the less appealing the testing + disclosure-to-partner option.

Consequently, if a sufficiently high fraction of the population gets tested while the rest does not (perhaps because it is too costly to do so), then we may face the scenario exposed by Kremer [1994] that having fewer partners may be consistent with increasing the long-run level of AIDS; except that here, some of the tested individuals have more exposures but fewer (actually zero) in the common pool. Even though welfare can be higher with disclosure, the equity issue is that the tested have a great advantage and the non tested lose enormously. If testing is more accessible for only certain categories of people – in this model, the old, but in practice it could be the rich – then it could become a major concern.

Similarly, this discussion could be reinterpreted by asking how much monogamy is needed to contain the disease and how more monogamous relationships influence welfare. Recall that age is only privately observable. Therefore, in the current model, monogamy between non-infected partners who can verify each other’s status, and assortative matching by negative test results are equivalent.
5.2 Differential Mortality

In the benchmark model, the fact that infected individuals die at the same rate as uninfected people tends to exaggerate the extent to which the pool of partners is worsened under testing. So this, *per se*, will tend to magnify any perverse effect. Progress in medicine tends to make infected people live longer and longer, but still not as long on average as uninfected persons; moreover, treatments are not commonly available in many countries. A simple way to reduce life expectancy in the model is to make those who were infected in the first period survive at the end of the first period with a probability $\alpha$.

When young infected people survive with a certain probability $\alpha$ at the end of the first period, the systems of equations (2) and (3) change in a relatively simple way.\(^{32}\)

It still turns out that practical prevalence $W$ is higher and welfare is lower under testing when one allows for a shorter expected lifespan among the infected (through different parametrizations for $\alpha$). However, the differences are dramatically smaller with small changes in $\alpha$: for example, for $\theta = 80$, and $\alpha = 0.8$, the difference in practical prevalence $W$ is only $0.06 - 0.02 = 0.04$, whereas it is $0.25 - 0.16 = 0.09$ in the benchmark case ($\alpha = 1$). The maximum value of $\alpha$ necessary to be consistent with eradication is the same in both cases and is given in Figure VI:

\(^{32}\)Calculations available from the author upon request.
Another way of interpreting the results is that while anti-retroviral therapies lower the differential mortality between HIV positives and negatives, and thus utility differences, ceteris paribus, they might have the adverse effect of increasing prevalence and reducing welfare overall. For a policy maker, this is a particularly delicate problem that cannot be realistically addressed within the simple framework of the current paper. I will nonetheless outline its key features below.

New treatments for AIDS reduce the cost of getting the disease, and therefore would tend to increase risky sexual behavior among susceptibles. Similarly, those drugs extend the lives of potentially contaminating individuals. On the other hand, while the prospect of HIV/AIDS as a manageable disease might decrease susceptibles’ appreciation of risk and therefore affect their behavior regarding safety, life-extending drugs for HIV patients are tantamount to an increase in HIV prevalence in the eyes of susceptibles. Thus, in a nutshell, knowing that more HIV-infected people are around because of the life-extending drugs, susceptibles might be induced to be more cautious, to decrease unsafe sex, which would therefore act as a counteracting force. Moreover, treatments and the steady counseling that goes with them might lead some HIV-positive people to decrease unsafe sex, and might also reduce the biological transmission factor, conditional on risky sexual behavior. Overall, the effect of new treatments on the spread of the epidemic is fundamentally ambiguous. Similarly, the possibility of being more effectively treated is expected to increase the demand for testing since only testing
opens access to treatments. A practical consequence is that changes in the observed proportion of positive tests have to be used with caution to infer changes in the overall risk level.  

5.3 ALTRUISM

Here I investigate how many altruists are needed to offset the potential negative effects of testing. I call altruistic the behavior of someone who after being informed of an HIV positive result, does not enter the general pool of contacts, and without any compensation: the altruist is assumed to derive the same utility from not entering the general pool as the selfish who enters it. I am not considering any policy promoting altruism because there is no meaningful way to assign a cost to it without data. Therefore, the percentages considered below have to be seen as exogenous.

Specifically, I ask three questions within the simple framework of this paper: how many altruists are needed in the steady-state with testing to keep practical prevalence (W) unchanged, compared to the steady-state without testing? Similarly, how many altruists are needed to keep welfare unchanged? How many altruists are needed to bring the steady-state with testing to a zero prevalence level (eradication)? The answers are summarized in the following figures: 

---

33 Geoffard and Mechoulan [2004] address this problem and provide some answers using San Francisco data.

34 The answers to the first two questions are virtually undistinguishable.
Percentage of "altruists" necessary to keep Practical Prevalence (W) and Welfare unchanged

Figure VII

Minimum Percentage of "altruists" necessary for eradication

Figure VIII
Interestingly, at the lowest levels of the parameter $\theta$ compatible with a strictly positive steady-state of prevalence in the benchmark case, only about 2.5% of altruists are needed for disease eradication. Note that for each steady-state prevalence level, it is possible to infer all possible combinations of percentages of altruists and numbers of partners compatible with that prevalence level. This in itself could prove to be valuable, for if we have an estimate of one, we can derive an estimate of the other.

To summarize, according to Figure VII, assuming the implementation of a testing program and a counseling policy ensuring a minimum fraction of approximately 14% altruists is not “too costly,” the conclusion that emerges is that testing will always be beneficial. In particular, the proportion of altruists that can be inferred from the public health studies is much higher. Similarly, for the range of $\theta$ corresponding to low or medium prevalence levels, Figure VIII tells us that the proportion of altruists necessary to eradicate the disease is smaller than what can be inferred from the public health literature. This result strongly supports the view that the disease survives because of the behavior of individuals who are unaware of their infection.

Consequently, in practice, assuming that the proportion of altruists is not significantly higher in the currently testing population than in the currently non-testing population, it appears that universal testing, or a mandatory testing program, would work effectively, not only in order to contain the disease, but also to eradicate it.

6 CONCLUSION

This paper has formally shown, in the simplest dynamic framework of HIV transmission, how to generate the Philipson-Posner conjecture on HIV testing. To summarize, in a base-case model with random mixing, everybody living for the same amount of time and behaving selfishly, testing would result in a higher steady-state prevalence and lower welfare. However, the main contribution of this work is to show that those results are fragile. Partial assortative matching and monogamy – when only a fraction of the population tests – have ambiguous effects that need to be better assessed. Introducing differential mortality in the model dramatically reduces the potential negative effects of testing. Most importantly, very little altruism seems needed for testing to become beneficial.

35 Even in this base-case model, a planner could theoretically use the test results for the benefit of society by organizing centralized monetary transfers from the non-infected to the infected. Such transfers would provide incentives for voluntary public disclosure so that the test information becomes a public good. This scheme could make both infected and non-infected better off. See Mechoulan [2001].
The public health literature actually tells us that a large proportion of individuals behave altruistically when tested positive and appropriately counseled. Beyond the mere availability of testing, the results further suggest combining counseling, the fostering of long-term relationships and education on prophylactic measures with universal or mandatory testing to help eradicate the disease.\(^{36}\)

The model presented is very stylized and can be refined along several dimensions. Possible extensions may include allowing for groups with different preferences to mix with one another, and for the infected to become infectious immediately. Above all, one may want to incorporate partners asking for protected sexual activity, heterogeneity in the cost of testing (making testing an endogenous decision), and treatment availability. Although this is bound to create severe complications, it is probably the most fruitful direction of research in open population models of HIV transmission.

**REFERENCES**


\(^{36}\) Further, recall that in many countries, getting tested *per se* carries a large social stigma. One of the additional benefits of having more people tested in such countries would be a significant reduction in the cost of stigma, which would make a universal testing program easier to implement, and hence create a compound welfare-improving effect.


