Sex and education model: Characterizing exposure across the HIV epidemic

by

Christian Aleman

Advisor

Raul Santaeulalia Llopis

Master project submitted in partial satisfaction of the requirements for the degree Master in Models and Methods of Quantitative Economics

Barcelona, September 20, 2018

IDEA-QEM

Abstract

In this paper I examine the role of education on the evolution of the exposure to HIV risk across the HIV epidemic. Cross country empirical results show that the education gradient in HIV is U-shaped. This means that at some point during the HIV epidemic, agents with higher education have a higher risk of becoming HIV positive. I propose a heterogeneous agent general equilibrium model to account for this stylized fact. Additionally, the paper proposes an algorithm to fully characterize the evolution of the HIV epidemic across stages. The model is calibrated to match the prevalences of the Malawian HIV epidemic. Finally, using the data generated by the model I calculate the education gradient for Malawi, the results show that the shape of the education gradient is in line with the empirical evidence. The model paves the road further policy implications, for example targeted prevention depending on the stage of the HIV epidemic.

Contents

List of Tables

List of Figures

1 Introduction

The negative consequences of the HIV virus range from individual impediments and the risk of death, to macroeconomic effects in detriment of economic growth. Since the discovery of the AIDS virus 35 years ago, several agencies started to work towards a common objective: complete eradication of the virus.

The challenge of erradication is larger in developing countries where the effectiveness of policy intervention is conditioned by financial constraints. However this is not the only hurtle. Governments struggle because they have limited understanding of the effects of policy intervention across the HIV epidemic. In particular the UNAIDS (2003) report, states that intervention should be based on specific epidemiology facts, such as the stage of the epidemic and those groups at high risk. This is recommended because policy efficacy might as well vary over the evolution of the epidemic.

Targeted prevention suggests that irrespectively of the stage of the epidemic an efficient way to reduce HIV diffusion is to target the groups at high risk. As described by the UNAIDS (2003) report, certain groups have a higher chance to get and spread HIV. These groups include most of the time individuals engaging in high-risk sexual behavior such as sex workers and their clients, injecting drug users and recently, men who have sex with men. How about individuals across educational groups?

Beegle and de Walque (2009) argue that there is a lack of consensus on the sign and the size of the relationship between education and HIV exposure, in other words, it is still not clear if people with lower education have a higher risk of infection than educated individuals. Santaeulalia-Llopis (2016) provides empirical evidence on the relationship between HIV status and education. He finds that risky behavior that increases HIV exposure, differs across educational groups as the HIV epidemic evolves. In particular, the education gradient in HIV has a U shape across the HIV epidemic. This means that in the early stages of the epidemic, additional years of education are associated with an increase of the probability of being HIV infected. This fact is the main motivation for this paper.

In this paper I propose a general equilibrium model that studies extramarital risky sex and its implications for HIV diffusion. The model treats the market of extramarital risky sex as any other competitive market, in which agents pay a competitive price and consume an amount of the desired good. In the model, agents differ principally in their education level and their level of asset holdings. Agents can be of two types: sex buyers or sex producers. The model has the feature that the probability of getting infected depends on the amount of extramarital risky sex an individual consumes. That is, the infection probability is endogenously determined by the model. In principle, agents are infinitely lived, however they have a probability of dying. Additionally, population grows at a constant rate.

The model divides the HIV epidemic in four stages, each with different features that intend to capture the particularities found in the data. In the first stage (*pre-epidemic stage*) there is no infection risk. Infection starts in stage two ((*myopic stage*)) where infected individuals suffer a productivity loss and therefore are payed less for their labor. Moreover, infected individuals have a higher probability of dying. However, in this stage agents are not aware of the risk of infection and they don't take into account the new characteristics of the environment to make their consumption choices. In other words, agents are myopic because they might be infected and observe the productivity loss, but they don't know the reason for this loss and they ignore it. In stage three (*maturity stage*), agents update their beliefs and take into account the probability of infection in their consumption decisions. Finally in stage four (*ARTs stage*), antiretroviral treatment is provided for all infected individuals. Introduction of antiretroviral drugs represent a gain in productivity, a reduction of the probability of infection and an increase in the survival probability of infected individuals.

I propose an algorithm that joins all four stages of the model with the intention to characterize the complete evolution of the HIV epidemic. Each stage after the first arrives as an unpredictable shock. The algorithm intends to capture the behavior of the agents in each stage, in particular the feature that during the *myopic stage* and *ARTs stage*, agents with higher education might have a higher exposure to HIV infection because of their decision to engage in more extramarital risky sex[1]. However, once they realize the negative effects of the disease and learn about the risk of infection, they decide to reduce their consumption of extramarital risky sex, and therefore reducing their probability of infection. In the *ARTs stage*, agents regain productivity levels even if they are infected, this in turn reverts the results of the *maturity stage* by making educated agents increase their risky sexual activities and therefore increasing HIV exposure. Further simulations of the model can be conducted to study the transitions between epidemic stages, however these simulations are left for future research.

Heterogeneous agent models are very rich in the sense that they generate data even without performing any type of simulation. I exploit this feature of the model

[1]However this might not be the case of Malawi

to extract the associated prevalences and match them to the different stages of the Malawian epidemic. It is in this way that the model is calibrated for Malawi. The data the model provides includes the prevalences across stages of the different types of agents and levels of education, proportion of infected individuals among the educated people and proportion of infected sex buyers and sex producers, relative to the total population. I use this information to run a linear probability model to explain HIV exposure using as explanatory variables, the level of education of the population and its interactions with each stage of the epidemic. In addition, I provide a disaggregation by gender to better compare the results to the data. This requires an additional assumption regarding the gender of the individuals. For simplicity reasons I assigned to men, the role of sex buyers and to women, the role of sex producers, however there is no reason that restricts contrary.

The results are not surprising as they are in line with the empirical results. Malawi doesn't have a clear U shaped educational gradient at the aggregate level, however the U shape is clearly visible for women. During the *myopic stage*of the HIV epidemic finishing secondary education increases the probability of infection by 2.9% among the total population. Among males and females 4.41% and 2.42 % respectively. During the *maturity* of the epidemic finishing secondary education reduces the chanced of infection by 1.45%; 0.6% for men and 1.8% for females. Finally during the *ARTs stage* education actually reduces the risk of infection by 3.16% among the total population and among men and women 3.39% and 0.79 % respectively.

The rest of the document is organized as follows, Section2 summarizes the state of the art regarding HIV diffusion and education. Section3 describes the evolution of the Malawian HIV/AIDS epidemic. Section4 makes a detail description of the proposed model. Section5 presents the solution algorithm, finally Section6 explores the results of the model and the calculation of the education gradient for Malawi, Section7 concludes. Further resources are provided in the AppendixA

2 Relevant literature

The literature on HIV exposure and education is very small. In fact, there is limited research that propose a theoretical framework to study the spread of HIV/AIDS, the use of general equilibrium models to asses implications of the HIV epidemic are on its infancy Greenwood et al. (2017). The present study intends to fills this gap.
An exeption to the previous claim is the work by Yao (2016). She constructs a life cycle model that relates sex decisions, fertility and education to HIV risk. Females choose between committed sex and casual sex. Committed sex tends to result in more premar-

ital children because premarital relations and committed sex facilitate marriage. When HIV risk increases, women choose to reduce premarital children and choose committed sex, since premarital sex and casual sex are positively related to HIV infection. Women with higher levels of education tend to have less committed sex and are more likely to choose casual sex, since having children has a higher opportunity cost for educated women. However, less educated women choose committed sex since they have a lower opportunity cost. The author proposes three policies to mitigate HIV exposure and concludes that an education subsidy is the most effective in reducing HIV prevalence, due to the reduction of premarital risky sex practices and the increase of the opportunity cost of child bearing.

Fortson (2008) also argues that education has a positive relation with premarital risky sex that increases the probability of HIV infection. Using data from the Demographic and Health Survey (DHS) he finds evidence of a positive education gradient in HIV infection. This means that high levels of education are more likely to get HIV infected. However this applies to individuals with very high levels of education, namely six years of schooling or more. Mishra et al. (2007) find a link between household wealth and HIV status, they document that in Sub Saharan Africa the wealthiest quintiles have a higher prevalence of HIV, and prevalence increases monotonically with wealth. This positive association can partly be lined to other factors such as geographical position of the household and the level of education.

Other authors use a different approach, and find opossite results. Case and Paxson (2013) study the role of education on the origins and geographic concentration of the HIV epidemic. He argues that the regions where HIV grew quickly, had an increase of non-marital sex activity and growing female education. Other autors like Preston (1996) and Bledsoe (1990) argue that education only reduces the chances of HIV infection if the health risks are completely understood. Additionally, education might increase the chance of poligamous partnerships and reduce early mariage that in turn increases the chance of infection.Oster (2012) tracks back the problem to an important increase of risky sexual behavior among educated women. Brent (2006) used a cross-section of 31 Sub Saharan countries to examine the effect of the level of education of females on the prevalence of HIV/AIDS, they find that female education is positively related with the infection rates. However when disentangling between primary and secondary education and accounting for different types of education within them [2], their results show that education is negatively related to female HIV infection if the

[2] for example consider all the age groups in the sector(gross) or as an alternative consider only those who are at the official school age (net)

variable *"females on the official school age"* is used instead of the alternative.

Following a similar argument, de Walque (2007) finds a negative relation between education and HIV diffusion. The author uses individual level data on the effectiveness of an information campaign to prevent HIV/AIDS in Uganda. They find that education is associated with a lower HIV infection risk among educated young females. They claim that educated women are more responsive to information campaigns. In the same line Alsan and Cutler (2013) find that girls enrollment in secondary education significantly increased sex abstinence, this in turn coincides with the dramatic fall of HIV prevalence in Uganda in the early 1990's.

In an attempt to reach a generalizable answer to the mixed evidence about the effect of education on HIV difussion,Santaeulalia-Llopis (2016) constructs an algorithm that characterizes the stages of the HIV epidemic across time and space, and computes the HIV education gradient for a panel of Sub Saharan countries. The author shows that the education gradient in HIV has a U shape over the course of the epidemic. Moreover men and women share the same U pattern but women have more pronounced shape. Specifically he concludes that in early stages, women with five additional years of education face a 7.4 % rise in the probability of infection. This in contrast to 3.8% in the case of men. In the middle of the epidemic this numbers reduce to zero for women and negative for men. In advanced stages the gradient becomes positive again with 2.7% for women and 1.75% for men[3]. This dynamic relation between the education level of the population, HIV prevalence and the stage of the epidemic is the starting point for the present research.

A richer model was proposed byGreenwood et al. (2017), where they analyze the Malawian HIV epidemic through the lenses of a choice theoretic general equilibrium search model. In their model, people choose among a set of sexual practices by being aware of the associated risk. They argue that the effectiveness of intervention depends also on a behavioral effect that is capable of having adverse macro effects by increasing HIV prevalence. This behavioral effect involves, engaging in more risky sexual practices, and lower incentive of protection. Kremer (1996)formally models partner choices in an attempt to introduce behavioral considerations into epidemiology, he argues that HIV risk is associated with the people's sexual activity choices. A higher risk causes individuals with a low sexual activity to reduce it even more, but high activity people reduce their activity less than the low activity ones, or even increase it. This differences

[3]Refer to the AppendixA for a visual presentation of the educational gradient found by Santaeulalia-Llopis (2016)

might result in an less favorable equilibrium for the economy.

Manuelli (2015) analyzes the effects of HIV/AIDS on aggregate output by studying the influence of the disease on investment directed to education and job training. He builds a human capital accumulation model, where individuals accumulate human capital in the form of schooling and job training. The accumulation decisions are influenced by the disease, since an infected individual might face a higher probability of permanent disability or death which in turn can significantly limit the individuals maximum level of effort, and at the macro level, cause a reduction of output per-capita.

3 Country Profile Malawi

"Malawi has one of the highest HIV prevalences in the world despite the impressive progress the country has made in controlling its HIV epidemic in recent years."Avert (2018)

Since the first case of HIV-AIDS diagnosed in Malawi, the country has suffered from a rapid increase of the number of people infected with the virus. Figure5 in the Appendix[4] shows the evolution of the Malawian HIV prevalence and Figure6 shows the evolution of the number of people living with HIV, new infections and mortality due to HIV. We can see from Figure5that the Malawian epidemic reaches its peach in year 1999 where the prevalence rate reached 16.7 percent of the population. After the peak the prevalence has been constantly decreasing, nevertheless still high compared to other countries. However, when looking at Figure6 we can see that the number of people living with HIV increases drastically until the year 2000, but after that it keeps growing until today. Although the share of infected people is reducing, the increase of the total number of infected could not be mitigated. On the other hand, the number of new infections of HIV/AIDS has been constantly reducing and it has reached its lowest point in 2016. We can also see that the number of deaths due to HIV/AIDS reduced since 1999, however it did not reach its lower value of 1990.

To better understand the Malawian epidemic it is important to dig deeper as transmission rates vary depending on gender, age and other socioeconomic variables such as the level of education, aspect that has not been deeply studied. Transmission happens mostly due to unprotected heterosexual sex between either cohabiting or married couples. In fact, premarital sex is common in Malawi to encourage marriage, although

[4]For presentation purposes, all of the figures in this section have been taken to the AppendixA.2

this is considered as a risky practice. Infection also happens due to mother to child transmission, and men who have sex with men.

Figure8 shows the share of females as a percentage of the total population living with HIV. It is alarming to see that more than half of the population who is infected are females. Moreover, this share has been constantly increasing, but since 2016 the rate started to grow faster reaching 60 percent of the infected population in 2016. Figure7 the number of AIDS related deaths by gender. Since the start of the epidemic the number of infected children increased by a factor of 3.5 up to 2003, it was only in 2004 that this growth stopped, and only in 2011 it started to reduce. A similar pattern can be spotted when looking that the number of deaths related to HIV, as only after 2004 the number of deaths started to reduce and almost reached pre-epidemic levels by 2016. It is important to note that women infected with HIV have a higher mortality rate than men until 2011, after that year, the situation reversed.

UNAIDS identified: women, young people engaged in early sexual activity, sex workers, gay men and other men who have sex with men as the most vulnerable sectors of the population in Malawi. But what about people with different education levels?. As strange as it may sound, it is not the less educated individuals who have a higher HIV prevalence, but exactly the opposite. Table1 shows the prevalence rate dis-aggregated by education level. Table2 shows the educational attainment in Malawi and its disaggregation by gender. We can see that the amount of people that finished secondary school is very small, in turn there is a very small number of females that finished high school. Although the situation slightly improved, Malawi still has a long way to go to increase educational attainment.

Table 1: Malawi, HIV prevalence among the general population (in %)

	Total	**By education level**	
		Primary or less	**Secondary or more**
DHS 2004	11.8	11.1	14.0
DHS 2010	10.7	10.2	11.8

Sources: The prevalences have been taken form DHS data.

Table 2: Malawi, Population over age 6 who completed secondary education(in %)

Survey	Share of (*) with secondary education or higher		
	(*)Total population	(*)Females	(Males)
DHS 1992	9.4	4.4	14.3
DHS 2000	16.0	11.1	20.9
DHS 2004	21.4	15.5	27.2
DHS 2010	25.6	20.0	31.2
DHS 2015	31.2	25.8	36.5

Sources: The prevalences have been taken form DHS data.

Table1 shows that the overall HIV prevalence in Malawi reduced from 11.8 percent in 2004 to 10.7 in 2010, the same occurs when we see the disaggregation by educational attainment. The prevalence of the people with primary or less education decreased slightly, keep in mind that this is a period in which the country expected the largest reduction in HIV prevalence at the national level. In the other hand, the prevalence among individuals with secondary education and higher, decreased more. However, when comparing the prevalence between less educated individuals and people with a higher educational attainment, we can see that every year, those who are more educated face a larger risk of getting infected. The gap reduced subtancialally in six years, nevertheless it still remains. Santaeulalia-Llopis (2016) documents this fact for a number of Sub Saharian countries and attributes the cause to the fact that educated individuals believe they are les vulnerable therefore they increase their risky sexual practices.Greenwood et al. (2017) studies the possibility of endogenous behavioral macro effects that play an important role when determining the effectiveness of intervention. They claim that up until a point these behavioral macro effects might reverse the positive effects of treatment.

The Malawian government, in cooperation with various international donors (UNAIDS, USAID to name the most important) implemented intensive mitigation programs during the 90's. These included the promotion of male circumcision and the treatment of other sexually transmitted diseases, but the most important and until now the most effective was the introduction of antiretroviral therapy in 2005 . By 2012 almost forty percent of the population living with HIV received antiretroviral treatment and by 2016 the threshold surpassed fifty percent, the malawian governement seeks to increase the level of coverage to 90 percent to achieve the UNAIDS 90,90,90[5]

[5]The 90,90,90 targets stand for: 90 percent of people living with HIV knowing their status, 90

HIV prevention goals. Although the Malawian governement claimed the introduction of antiretroviral terapy was a complete success due to the aparent reduction of HIV prevalence, researchers suggest the effect is more ambiguous than expected Greenwood et al. (2017). Despite all the efforts Malawi remains with one of the highest HIV prevalence rate in the world.

4 The model

This section describes the proposed heterogeneous agent general equilibrium model.

4.1 Model characteristics

The model features two types agents: sex consumers and sex producers. They are finitely lived, have rational expectations and intend to maximize future discounted utility, $E_0 \sum_t^\infty \beta^t u$, subject to their respective budget constraints. I assume the function $u : [0, \infty) \to \mathbb{R}$ is strictly increasing, strictly concave and twice continuously differentiable.
Agents interact in three competitive markets in this economy: The goods market, the sex market and the assets market.

- From the goods market agents can obtain consumption good c at price p_1[6].
- Non-marital risky sex x is traded in the sex market for price $p_2 = p$. Where variable x is continuous.
- Individuals can either save or borrow in the asset market by trading a non-contingent asset a with endogenous return r.
- Agents education can either be high or low $e = \{1, 0\}$.
- HIV status can be positive or negative $h = \{1, 0\}$.
- Labor is supplied inelastically.
- Agents receive labor income $y(e)$, where income differs according to their education level e. In particular $y(1) > y(0)$.
- Let γ_e denote the survival probability of an agent with level of education e.

percent of people living with HIV who know their status are on treatment and finally 90 percent of people on treatment are viraly supresssed

[6]From now on I will use the price of the consumption good as numeraire, then $p_1 = 1$

- Let ψ^h be the fertility rate of an individual with HIV status h.
- Let $\lambda(x, h)$ denote the probability of getting infected with HIV. Note that the probability of infection is actually a function of the amount of risky sex(x) consumed by the agent. This value is determined endogenously.
- There are two types of agents sex buyers (g) and sex producers ($-g$).

Following the discussion in Section1, the model intends to characterize the different stages of the HIV epidemic according to the level of education of the population. In particular, the increase in the probability of infection for the most educated in the early and final stages.
In line with Santaeulalia-Llopis (2016), four stages of the epidemic have been identified.

1. Pre-Epidemic stage
2. Miopic stage of the epidemic
3. Difussion and maturity of the epidemic
4. Introduction of anti-retro viral drugs

The features of each stage will be described in detail in the upcoming sections.

The model features additional dynamics, in the sense that it intends to capture the evolution of the HIV epidemic starting from stage one until stage four. For presentation purposes each stage characterizes a stationary equilibrium, but later on they will be linked with the intention to describe the complete evolution of the HIV epidemic.

In the following sections I describe in detail the characteristics of each of the stages of the model.

4.2 Pre-epidemic stage

In this pre-HIV world, there is no probability of infection, namely $\lambda_(x, h) = 0$. Agents differ on their type g, assets a and education level e.

Sex Buyers:
For agents of type g the dynamic problem is:

$$V(a, e, g; \Phi) = \max_{c \geq 0, x \geq 0, a' \geq 0} u(c, x) + \beta \gamma_e V(a', e, g; \Phi') \tag{1}$$

$$\text{s.t}$$

$$c + px + a' = y(e) + (1 + r(\Phi))a \tag{2}$$

Agents interact in non-marital risky sex activities, these agents buys non-marital risky sex x at price p and saves or borrows a' with return r. Note that labor is supplied inelastically and depends on their level of education e, then their labor income is $y(e)$, where $y(1) > y(0)$.
This problem is characterized by three individual state variables (a, e, g) and one aggregate state variable Φ, which represents the population distribution over the states a, e, g. The choice variables are c, x and a'.
The above problem can be written in sequential form:

$$\max_{c_t, x_t, a_{t+1}} E_0 \sum_{t=0}^{\infty} \beta^t u(c_t, x_t)$$
$$\text{s.t}$$
$$c_t + p_t x_t + a_{t+1} = y(e) + (1 + r_t) a_t$$

FOC's:

$$\frac{\partial u(c_t, x_t)}{\partial x_t} = -u'_c(c_t, x_t) p_t + u'_x(c_t, x_t) = 0$$
$$\frac{\partial u(c_t, x_t)}{\partial a_{t+1}} = -u'_c(c_t, x_t) + u'_c(c_{t+1}, x_{t+1})(1 + r_t)\beta = 0$$

Given the price system $\{p_t\}_{t=0}^{\infty}, \{r_t\}_{t=0}^{\infty}$ the solution is characterized by the sequence of allocations $\{c_t\}_{t=0}^{\infty}, \{x_t\}_{t=0}^{\infty}, \{a_t\}_{t=0}^{\infty}$ such that they solve the above maximization problem.

Sex Producers:
For agents of type $-g$ the dynamic problem is:

$$V(a, e, -g; \Phi) = \max_{c \geq 0, 1 \geq l \geq 0, a' \geq 0} u(c) + \beta \gamma_e V(a', e, -g; \Phi') \tag{3}$$
$$\text{s.t}$$
$$c + a' = p l^{\alpha} + y(e)(1 - l) + (1 + r(\Phi)) a \tag{4}$$

Note that for these agents, extra marital risky sex does not generate any utility, then the utility function only depends on the amount of c consumed. Additionally, these agents produce sex with a decreasing return to scale production function $x = l^{\alpha}$ where $\alpha \in (0, 1)$. In other words, these agents produce non-marital risky-sex x using time l and sell it a price p. Where l is the fraction of labor dedicated to the production of sex, the rest of the labor $(1 - l)$ which is not allocated to sex production is sold in the market for labor income $y(e)$. Additionally agents save or borrow assets a at rate r.

The above problem can be written in sequential form:

$$\max_{c_t, l_t, a_{t+1}} E_0 \sum_{t=0}^{\infty} \beta^t u(c_t)$$

$$\text{s.t}$$

$$c_t + a_{t+1} = p_t l_t^{\alpha} + y(e)(1 - l_t) + (1 + r_t)a_t$$

FOC's:

$$\frac{\partial u(c_t)}{\partial l_t} = -u_c'(c_t)(\alpha p_t u^{\alpha-1} - y(e)) = 0$$
$$\frac{\partial u(c_t)}{\partial a_{t+1}} = -u_c'(c_t) + u_c'(c_{t+1})(1 + r_t)\beta = 0$$

Given the sequence of prices $\{p_t\}_{t=0}^{\infty}, \{r_t\}_{t=0}^{\infty}$ the solution are the sequences of allocations $\{c_t\}_{t=0}^{\infty}, \{l_t\}_{t=0}^{\infty}, \{l_t\}_{t=0}^{\infty}, \{a_t\}_{t=0}^{\infty}$ that solve the agent's problem.
For the computation of the solution I assume that the utility function is CRRA with relative risk aversion parameter σ, all agents regardless of their type share the same relative risk aversion. Then $u_c'(c_t, x_t) = c_t^{-\sigma}$, $u_x'(c_t, x_t) = x_t^{-\sigma}$.

Solution to the recursive problem

Given prices p, r the solution to the recursive problem of agents g and $-g$ are the policy functions $a(a, e, g; \Phi), x(a, e, g; \Phi), c(a, e, g; \Phi), l(a, e, g; \Phi)$ that induce a stationary distribution $\Phi(\mathcal{A}, \mathcal{E}, \mathcal{G})$ over the set of state variables. Denote Φ as the aggregate state variable.

The aggregate sate variable and transition function

The aggregate state variable evolves according to:

$$\Phi' = F(\Phi) \tag{5}$$

Where the function $F : \mathcal{M} \rightarrow \mathcal{M}$ is the aggregate law of motion, mapping distributions to distributions. F summarizes how agents move within the distribution of assets , education and type from one period to the next, however this is exactly what a transition function tell us.

Define the transition function $\mathcal{Q} : \mathcal{Z} \times \mathcal{B}(\mathcal{Z}) \to [0,1]$ by:

$$Q((a,e,g)(\mathcal{A},\mathcal{E},\mathcal{G})) = \begin{cases} \gamma_e & \text{if} \quad a(a,e,g;\Phi) \in \mathcal{A} \\ 0 & \text{else} \end{cases}$$
$$\forall \ (a,e,g) \in \mathcal{Z} \ \text{ and } \ (\mathcal{A},\mathcal{E},\mathcal{G}) \in \mathcal{B}(\mathcal{Z})$$

Where $\mathcal{Z}$ consists of all n-tuples of $A \times E \times G$.
Define $\mathcal{B}(\mathcal{Z})$ as the set of Borel sets on $\mathcal{Z}$, in particular $\mathcal{A},\mathcal{E},\mathcal{G} \in \mathcal{B}(\mathcal{Z})$ where $\mathcal{A},\mathcal{E},\mathcal{G}$ are projections of $\mathcal{Z}$ over the spaces A, E and G respectively. Let $\mathcal{P}$ be a probability measure on $\mathcal{B}(\mathcal{Z})$, then $\mathcal{P} : \mathcal{B}(\mathcal{Z}) \to [0,1]$.
Then the evolution of the asset distribution is:

$$\Phi'(\mathcal{A},\mathcal{E},\mathcal{G}) = F(\Phi)(\mathcal{A},\mathcal{E},\mathcal{G}) = \int_{a,e,g} Q((a,e,g)(\mathcal{A},\mathcal{E},\mathcal{G}))d\Phi + \psi\Phi((a',e,g)(\mathcal{A},\mathcal{E},\mathcal{G})) \tag{6}$$

Which is the fraction of people with assets in $\mathcal{A}$, education $\mathcal{E}$ and type $\mathcal{G}$ as measured by Φ, that transits to $(\mathcal{A},\mathcal{E},\mathcal{G})$ as measured by $\mathcal{Q}$. The last term accounts for the new born. Population increases according to the fertility rate ψ, it important to note that individuals of a certain type give birth to people of the same type.

Stationary equilibrium

The *stationary equilibrium of the Pre-Epicemic stage* is:

- An interest rate r and price p
- Policy functions $a(a,e,g;\Phi), x(a,e,g;\Phi), c(a,e,g;\Phi), l(a,e,g;\Phi)$
- A stationary distribution $\Phi(\mathcal{A},\mathcal{E},\mathcal{G})$

Such that:

a Given r and p the policy functions $a(a,e,g;\Phi), x(a,e,g;\Phi), c(a,e,g;\Phi), l(a,e,g;\Phi)$ solve the sex buyers and sex producers problem respectively.

b The stationary probability distribution $\Phi'(\mathcal{A},\mathcal{E},\mathcal{G})$ is induced by the optimal policy $a(a,e,g;\Phi)$.

c All markets clear.

$$\int_{a,e,g} a(a,e,g;\Phi)d\Phi = 0$$
$$\int_{a,e,g} x(a,e,g;\Phi)d\Phi = \int_{a,e,-g} x(a,e,-g;\Phi)d\Phi$$

That is, there is zero net supply of assets, the sex markets clear and the consumption market clears by Walras law.

4.3 Miopic stage of the epidemic

The myopic stage is almost identical to the pre-HIV stage except for the following aspects:

1. Now there is the possibility of becoming HIV infected as the result of non-marital risky sex activity x. The probability of infection is $\lambda(x, h)$[7].

2. Agents are myopic. This implies that they are being infected but they are not aware of it, in fact they believe that at all periods $h = h' = 0$, i.e., that the probability of infection $\lambda(x, h)$ is zero. In particular, in this stage they naively set $V(a', e, g, h' = 0)$ when solving their maximization problem; it is in that sense that they are myopic.

3. In this stage the survival probability γ_e and labor income $y(e)$ not only depend on education e but also on the HIV status, h. Then $\gamma_e(h)$ and $y(e, h)$. In particular given e, $\gamma_e(1) < \gamma_e(0)$ and $y(e, 1) < y(e, 0)$.

Following Bethencourt and Rios-Rull (2009),the functional form that determines to which extent the amount of non-marital risky sex affects the probability of infection is:

$$\lambda(x) = \frac{exp(x)}{exp(x) + \rho exp(-x)}$$

The formula above depends only on one parameter, $\rho \geq 0$ and since $x \geq 0$ the domain of the function is, $\mathbb{R}_+$ and the range $\{\frac{1}{1+\rho} \leq \lambda(x) \leq 1\}$. We can interpret ρ as the parameter that captures the individual's degree of understanding of the epidemic evolution. More over, a lower ρ is associated with a lower degree of understanding, and a higher value of ρ means the agent is better informed about the risks of infection. Then $\frac{\partial \lambda(x)}{\partial x} > 0$ and $\frac{\partial \lambda(x)}{\partial \rho} < 0$ this is, the higher the understanding parameter ρ, the less the chances of infection.

From one side, even if individuals have no extra marital risky-sex $x = 0$ they still have chance $\frac{1}{1+\rho}$ of getting infected. From the other side, as x tends to infinity $\lambda(x)$ tends to one.

$$\lim_{x \to \infty} \lambda(x) = 1$$

[7] For simplicity I denote the of getting HIV infected given that an individual is healthy $\lambda(x, 0)$ by just $\lambda(x)$

Sex buyers

For agents of type g the dynamic problem is:

$$V(a,e,g,h;\Phi) = \max_{c,x,a'} \quad u(c,x) + \beta \sum_{h'|h} \gamma_e(h')\lambda(x,h'|h)V(a',e,g,h'=0;\Phi)$$

$$\text{s.t}$$

$$c + px + a' = y(e,h) + (1+r(\Phi))a$$

The FOC's and solution to the sequential formulation are analogous to the previous stage.

Sex Producers

For agents of type $-g$ the dynamic problem is:

$$V(a,e,-g,h;\Phi) = \max_{c,l,a'} \quad u(c) + \beta \sum_{h'|h} \gamma_e(h')\lambda(x,h'|h)V(a',e,-g,h'=0;\Phi)$$

$$\text{s.t}$$

$$c + a' = pl^{\alpha} + y(e,h)(1-l) + (1+r(\Phi))a$$

The FOC's and solution to the sequential formulation are analogous to the previous stage.

Solution to the recursive problem

Given prices p, r the solution to the recursive problem of agents g and $-g$ are the policy functions $a(a,e,g,h;\Phi), x(a,e,g,h;\Phi), c(a,e,g,h;\Phi), l(a,e,g,h;\Phi)$ that induce a stationary distribution $\Phi(\mathcal{A},\mathcal{E},\mathcal{G},\mathcal{H})$ over the set of state variables. Φ now includes a new state variable, the HIV status h.

The aggregate sate variable and transition function

As before the aggregate state variable evolves according to:

$$\Phi' = F(\Phi) \tag{7}$$

Where $F : \mathcal{M} \to \mathcal{M}$ now summarizes how agents move within the distribution of assets , education, type, and HIV status from one period to the next.

Define the transition function $Q : \mathcal{Z} \times \mathcal{B}(\mathcal{Z}) \rightarrow [0,1]$ by:

$$Q((a,e,g,h)(\mathcal{A},\mathcal{E},\mathcal{G},\mathcal{H})) = \begin{cases} \gamma_e(h')\lambda(x,h'|h) & \text{if} \quad a(a,e,g,h;\Phi) \in \mathcal{A} \\ 0 & \text{else} \end{cases}$$
$$\forall \ (a,e,g,h) \in \mathcal{Z} \ \text{and} \ (\mathcal{A},\mathcal{E},\mathcal{G},\mathcal{H}) \in \mathcal{B}(\mathcal{Z})$$

Again $\mathcal{Z}$ consists of all n-tuples of $A \times E \times G \times H$, and $\mathcal{B}(\mathcal{Z})$ is the set of Borel sets on $\mathcal{Z}$. Moreover, $\mathcal{A},\mathcal{E},\mathcal{G},\mathcal{H} \in \mathcal{B}(\mathcal{Z})$ and $\mathcal{A},\mathcal{E},\mathcal{G},\mathcal{H}$ are projections of $\mathcal{Z}$ over the spaces A, E, H and H respectively. Let $\mathcal{P}$ be a probability measure on $\mathcal{B}(\mathcal{Z})$, then $\mathcal{P} : \mathcal{B}(\mathcal{Z}) \rightarrow [0,1]$.
Then the evolution of the asset distribution is:

$$\Phi'(\mathcal{A},\mathcal{E},\mathcal{G},\mathcal{H}) = F(\Phi)(\mathcal{A},\mathcal{E},\mathcal{G},\mathcal{H}) = \int_{a,e,g,h} Q((a,e,g,h)(\mathcal{A},\mathcal{E},\mathcal{G},\mathcal{H}))d\Phi + \psi^{h'}\Phi((a',e,g,h')(\mathcal{A},\mathcal{E},\mathcal{G},\mathcal{H})) \tag{8}$$

Which is the fraction of people with assets in $\mathcal{A}$, education $\mathcal{E}$,type $\mathcal{G}$ and HIV status $\mathcal{H}$ as measured by Φ, that transits to $(\mathcal{A},\mathcal{E},\mathcal{G},\mathcal{H})$ as measured by $\mathcal{Q}$. The fertility rate now depends on the HIV status of the individual. For simplicity I assume that $\psi^{h'} = \psi^h = \psi$, that is there is no difference between the fertility rate of the people infected with HIV and the non infected.

I now derive the evolution of the asset distribution explicitly for the case in which $h = \{0,1\}$ and $e = \{0,1\}$. Let:

$$\Gamma = \begin{bmatrix} \gamma_e(0) & 0 \\ 0 & \gamma_e(1) \end{bmatrix} \Lambda = \begin{bmatrix} 1-\lambda(x) & 0 \\ \lambda(x) & 1 \end{bmatrix} \Psi = \begin{bmatrix} \hat{\psi}^0 & 0 \\ 0 & \hat{\psi}^1 \end{bmatrix}$$

Where matrix Γ collects the survival probabilities of both types. The transition matrix Λ contains the probabilities of infection, note that once infected it is impossible to be cured, then $\Lambda_{1,1} = 1 - \lambda(x)$ is the probability of staying healthy, $\Lambda_{2,1} = \lambda(x)$ is the probability of turning HIV positive. Matrix Ψ collects the fertility rates, where $\hat{\psi}^h = 1 + \psi^h$, as before I assume $\psi^1 = \psi^0 = \psi$.
Then the evolution of the population with education e, before choosing a follows:

$$\begin{bmatrix} \phi_{t+1}(a_t,e,g,0) \\ \phi_{t+1}(a_t,e,g,1) \end{bmatrix} = \Psi \times \Lambda \times \Gamma \times \begin{bmatrix} \phi_t(a_t,e,g,0) \\ \phi_t(a_t,e,g,1) \end{bmatrix}$$

Define:

$$\Lambda \times \Gamma = \begin{bmatrix} (1-\lambda(x))\gamma_e(0) & 0 \\ \lambda(x)\gamma_e(0) & \gamma_e(1) \end{bmatrix} = \Omega$$

Then, together with the decision rule $a(a, e, g, h, \Phi)$ the transition matrix $\mathcal{Q}((a, e, g, h)(\mathcal{A}, \mathcal{E}, \mathcal{G}, \mathcal{H}))$is constructed as follows:

$$\mathcal{Q}((a, e, g, h)(\mathcal{A}, \mathcal{E}, \mathcal{G}, \mathcal{H})) = \sum_{h' \in \mathcal{H}} \mathbf{1}_{a \in \mathcal{A}} \Omega(h'|h)$$

And the endogenous asset distribution can be rewritten as:

$$\begin{aligned} \boldsymbol{\phi}_{t+1}(a_{t+1}, e, g, h_{t+1}) = \int_{a,e,g,h} \sum_{h_{t+1}|h_t} \mathbf{1}_{a_{t+1}=a(a,e,g,h)} \gamma_e(h_{t+1}) \lambda_e(h_{t+1}|h_t) d\boldsymbol{\phi}_t(a, e, g, h) \\ + \psi \boldsymbol{\phi}_t(a_{t+1}, e, g, h_{t+1}) \end{aligned} \tag{9}$$

This is equivalent to Equation8.

Stationary equilibrium

The *stationary equilibrium of the myopic stage* is:

- An interest rate r and price p
- Policy functions $a(a, e, g, h; \Phi), x(a, e, g, h; \Phi), c(a, e, g, h; \Phi), l(a, e, g, h; \Phi)$
- A stationary distribution $\Phi(\mathcal{A}, \mathcal{E}, \mathcal{G}, \mathcal{H})$

Such that:

a Given r and p the policy functions $a(a, e, g, h; \Phi), x(a, e, g, h; \Phi), c(a, e, g, h; \Phi), l(a, e, g, h; \Phi)$ solve the sex buyers and sex producers problem respectively.

b The stationary probability distribution $\Phi'(\mathcal{A}, \mathcal{E}, \mathcal{G}, \mathcal{H})$ is induced by the optimal policy $a(a, e, g, h; \Phi)$.

c All markets clear.

$$\int_{a,e,g,h} a(a, e, g, h; \Phi) d\Phi = 0$$
$$\int_{a,e,g,h} x(a, e, g, h; \Phi) d\Phi = \int_{a,e,-g,h} x(a, e, -g, h; \Phi) d\Phi$$

With zero net supply of assets. The sex markets clears and the consumption market clears by Walras law.

4.4 Difussion and maturity of the epidemic

This stage is similar to the previous one apart from two major differences:

- During the maturity of the epidemic agents now recognize that there is a probability of infection due to non-marital risky sex. This probability of infection is more accurately understood by those who are more educated $e = 1$. This means that ρ now depends on the the education level of the individual ρ_e, moreover $\rho_1 > \rho_0$.
- The agents are not myopic anymore. That is, they assign the correct continuation value when solving their maximization problem.

Sex buyers
For agents of type g the dynamic problem is:

$$V(a, e, g, h; \Phi) = \max_{c,x,a'} \quad u(c, x) + \beta \sum_{h'|h} \gamma_e(h')\lambda(x, h'|h)V(a', e, g, h'; \Phi)$$

$$\text{s.t}$$

$$c + px + a' = y(e, h) + (1 + r(\Phi))a$$

Sex Producers
For agents of type $-g$ the dynamic problem is:

$$V(a, e, -g, h; \Phi) = \max_{c,l,a'} \quad u(c) + \beta \sum_{h'|h} \gamma_e(h')\lambda(x, h'|h)V(a', e, -g, h'; \Phi)$$

$$\text{s.t}$$

$$c + a' = pl^{\alpha} + y(e, h)(1 - l) + (1 + r(\Phi))a$$

The FOC's and solution to the sequential formulation are analogous to the previous stage.
The construction of the transition function and the stationary distribution is identical to the previous stage.
The definition of the *stationary equilibrium of the maturity of the epidemic* is the same as the stationary equilibrium of the previous stage.

4.5 Antiretroviral treatment (ART's)

In this stage:

- There is anti-retroviral (ART) treatment $d = \{0, 1\}$ for everyone that is infected, this means that $h = 1 \Rightarrow d = 1$.

- ART drugs change affect survival rates $\gamma_e(h,d)$, probabilities of infection $\lambda(x,h,d)$ and individual productivity $y(e,h,d)$ of those who are infected $h=1$, there is no difference on the benefits of the drug among types g of agents.
  - Then the new survival rate of the infected $\gamma_e(1,1)$ is larger than before.
  - Also $y(e,0,0) > y(e,1,1)\ \forall e \in E$.

Sex buyers

For agents of type g the dynamic problem is:

$$V(a,e,g,h;\Phi) = \max_{c,x,a'} \quad u(c,x) + \beta \sum_{h'|h} \gamma_e(h',d)\lambda(x,h'|h,d)V(a',e,g,h';\Phi)$$

$$\text{s.t}$$

$$c + px + a' = y(e,h,d) + (1 + r(\Phi))a$$

Sex Producers

For agents of type $-g$ the dynamic problem is:

$$V(a,e,-g,h;\Phi) = \max_{c,l,a'} \quad u(c) + \beta \sum_{h'|h} \gamma_e(h',d)\lambda(x,h'|h,d)V(a',e,-g,h';\Phi)$$

$$\text{s.t}$$

$$c + a' = pl^{\alpha} + y(e,h,d)(1-l) + (1 + r(\Phi))a$$

The FOC's and solution to the sequential formulation are analogous to the previous stage.

The construction of the transition function and the stationary distribution is identical to the previous stage.

The definition of the *stationary equilibrium of the last stage of the epidemic* is the same as the stationary equilibrium of the previous stage.

An algorithm to find the stationary equilibrium of each of the epidemic stages is provided in the AppendixA.1.

5 Solution algorithm, keeping track of the HIV epidemic

The following algorithm intends to join the stages $1-4$ to provide a full dynamic characterization of the HIV epidemic. For the following algorithm, define the following stages according to the description in Section1:

Stage 1 (*Pre-Epidemic*): There is no probability of infection $\lambda(x,h)=0$. Agents solve the maximization problem described in Section4.2.

Stage 2 (*Miopic onset*): Agents are faced with a positive infection rate $\lambda(x)>0$, however they are not aware of it. Denote $\hat{\phi}(a,e,g,h=1)$ the initial distribution of infected individuals. The agents solve the maximization problem explained in Section4.3

Stage 3 (*Maturity*): Agents acknowledge the risk of HIV infection, and correct their beliefs accounting for the probability of infection. They solve the problem described in Section4.4.

Stage 4 (*ART's*): All infected individuals receive ART treatment. They face the maximization problem in Section4.5

Algorithm

Initialization

Step 1: Find the stationary equilibrium for the *pre-epidemic stage*[8].

Step 2: *The Myopic stage* arrives as an unexpected shock that hits the steady state of the *pre-epidemic stage*. However agents are not aware of it. Solve forward using the decision rules of the*pre-epidemic stage* but with the productivity levels, infection probabilities and survival rates of the *myopic stage*.

Step 3: The *maturity of the epidemic* arrives as an unexpected shock that hits the transition of the *myopic stage* after t periods. This means that the *myopic stage* never arrives to the steady state because the economy gets hit by the *maturity* during the transition. Find the steady state of the *maturity stage*, start solving backwards for the proportion μ of agents that fully understand the probability of infection. Solve forward for those agents $1-\mu$ that are not aware of the risk. At some point, all agents will be aware of the risk.

Step 4: The introduction of *anti retro viral treatment ART* arrives as an unexpected shock that hits the transition after $\hat{t}$ periods of *maturity*. Again *maturity* never reaches the steady state. First compute the steady state of the *ARTs stage* and solve backwards until the turning point with *maturity* is reached.

5.1 Calibration

The model parameters were calibrated to match the prevalence across the ongoing Malawian epidemic. Note that there is no particular convention to identify in the data, any of the four theoretical stages described in the model. For example the *pre-epidemic stage* does not exist in real life, so there is no way to assign values to it. In addition the model does not discriminate between male and female individuals; disaggregation that is important when characterizing the education gradient by gender.

For calibration purposes I make an assumption regarding the gender of the individuals. I assign the role of sex buyers to men and sex producers to women. As described in Section1 there is no convention that prevents the types to be classified in the opposite way. From now on I will refer to type one individuals as men and type

[8] For an algorithm to compute the stationary equilibrium refer to the AppendixA.1

two as women. As it will be shows this assumption does not affect the main results of the paper.

I use data from 2000 for the calibration of the *myopic stage* of the epidemic, I chose 2000 since the Malawian HIV/AIDS epidemic reached a peak around that year. The HIV prevalence in Malawi in 2000 reached 16.8 percent of the total population, and this is the number I will be targeting. Unfortunately there is no additional data concerning the prevalence across educational groups by gender in 2000, so it is not possible to calibrate this numbers to values found in the data. However using linear interpolation methods by taking the prevalence values for 2004 as a reference, the estimated HIV prevalence for educated females reaches 19.3 percent in 2000, I therefore use this value as a benchmark to calibrate the endogenous prevalence for educated females that the model generates. [9]

The *maturity and diffusion* of the epidemic is calibrated to match the prevalence of the year 2004, in addition the actual value of the HIV prevalence rate among educated females and males will be used for the calibration. Finally, I use the data for 2010 to calibrate the *ART's stage* of the epidemic . Although antiretroviral treatment was introduced in Malawi in the year 2005, the effects are visible 5-6 years after, then it is reasonable to think 2010 as the last stage of the epidemic described by the model. Table3 through Table5 show the parameters inherent of each stage of the Malawian HIV/AIDS epidemic. The rest of the parameters which I assume are not stage specific are summarized in 6:

Table 3: Targeted Prevalences (in %)

Stage	**Year**	**Total**	**HIV/AIDS Prevalence of (*) who completed higher education or more**	
			(*)**Females**	(*)**Males**
Myopic Stage	2000	15.0	19.3	16.5
Maturity	2004	11.8	15.1	12.9
ARTs	2010	10.7	16.1	8.9

Sources: Taken form DHS data.

[9]Again, there is no particular reason other than the calculation of the HIV education gradient, for which I choose to match these prevalences .

Table 4: **Fertility and mortality rates**

Stage	**Year**	**Fertility rate(%)**	**Mortality rate including HIV/AIDS infected(%)**	**Mortality rate excluding HIV/AIDS infected(%)**
Myopic Stage	2000	6.20	1.68	1.35
Maturity	2004	5.95	1.45	1.04
ARTs	2010	5.30	0.98	0.71

Sources: World Bank, Sustainable Development Indicators.

Table 5: **Share of people educated but infected**

Stage	**Year**	**Share of (*) with secondary education or more(%)**		**Share of (**)with secondary education or more and are infected (%).**	
		(*)Females	**(*)Males**	**(**)Females**	**(**)Males**
Myopic Stage	2000	11.1	20.9	2.1	3.4
Maturity	2004	15.5	27.2	2.3	3.5
ARTs	2010	20.0	31.2	3.2	2.8

Sources: DHS data .

Table 6: **Parameters which are not stage specific**

Parameter Symbol	**Value**	**Description**	**Source**
μ	3.00	Risk Aversion	Standard value
β	0.96	Subjective discount factor	Standard value
α	0.1	Labor income share	-

6 Results

Once the model was calibrated in each stage respectively, the value functions an policy functions were approximated via value function iteration. Figure9 to Figure16 show the obtained value functions, and the policy functions for asset holdings, consumption and extra-marital risky sex consumption, by type of agent in each stage of the epidemic. Starting from stage 1 we can see that educated individuals hold a larger amount of assets than the less educated individuals, regardless of their type. Also from stage 2 we see that educated individuals who are not infected are wealthier, consume more and incur in more extra-marital risky sex than the less educated counterparts. Note that the effect is reversed when looking at sex producers. As expected, the policy function for sex production is flat, since the individual supplies a constant account of labor to the sex market.

A key result of the model is the endogenous reduction of extra-marital sex consumption once the risk for HIV infection is acknowledged. We can see this phenomenon by comparing Figures 12a,12b with Figures 16a,16a. These figures show the extra-marital risky sex consumption policy function for Stage 2 and Stage 3 respectively, it is clear than individuals realize that the chances of infection increase with the amount of extra-marital risky sex consumed, then they automatically decide to reduce the amount consumed.

Without the need to run simulations, we can take the data provided by the model and calculate the education gradient for the total Malawian population and its disaggregation by gender. The model outputs data in a panel structure. Since the model assumes a continuum of individuals I choose to generate a sample of size $N = 20000$. Only in this sense is this data more convenient than survey data. However, Survey data is richer in all senses, as it allow us to control for other aspects that the model cannot simulate. The data used to calculate the educational gradient includes: HIV prevalence of the whole population, and the portion of the population who is educated but infected, and its respective gender equivalents.

6.1 A linear probability model

Following Santaeulalia-Llopis (2016), I propose a linear probability model with the HIV status of the individual across stages as dependent variable and the level of education as an independent variable.

Consider the panel structure generated by the model $X_{i,t}$ where $y_{i,t}, s_{i,t} \in X_{i,t}$. Where i indexes the individuals $i \in \{1, ..., N\}$ and t indexes the stage of the epidemic $t \in \{2, 3, 4\}$[10]. Here $y_{i,t}$ denotes the HIV status of an individual i living in stage of the epidemic t; $edu_{i,t}$ denotes the education status of the individual i in stage t. Then the

[10]Note that t starts in Stage 2, since as explained in Section5.1 Stage 1 does not exist in real life

linear projection of y on s is:

$$y_{i,t} = \alpha_2 + \sum_{t>2} \alpha_t \mathbf{1}_t + (\gamma_2 + \sum_{t>2} \gamma_t \mathbf{1}_t) edu_{i,t} + \epsilon_{i,t} \tag{10}$$

$\mathbf{1}_t$ is the indicator function that takes the value of 1 if we are in epidemic stage t otherwise is zero. Given the structure of the model we can be sure that $E(\epsilon) = 0$ and $Var(\epsilon) = \sigma^2$, then the model can be estimated via OLS.

Keep in mind that in this model $y_{i,t}$ and $s_{i,t}$ are both categorical variables since the model only distinguishes between the individuals who finished secondary education and the ones who didn't. Given this, the calculation of the gradient differs from the conventional one. In this model the base category α_2 represents those individuals who are in the *myopic stage* and have not finished secondary education. γ_2 is the HIV-education gradient in the *myopic stage* of the epidemic. In the same way $\gamma_2 + \alpha_t + \gamma_t$ is the HIV-education gradient in stage t, then the difference in the HIV education gradient between an individual in stage t and the *myopic stage* is $\alpha_t + \gamma_t$. The interpretation of the HIV-economic gradient is as follows; in the*myopic stage*, finishing secondary education or getting higher levels of education changes the probability of being infected by γ_2 percent. In later stages, the probability of of being infected changes by $\gamma_2 + \alpha_t + \gamma_t$ percent.

Table7 shows the results of the linear probability model. This includes the models for the total population, and the dis-aggregation by gender.

Table 7: The HIV-education gradient for Malawi

	Dependent variable: HIV Status		
	Total population	Males	Females
Education	0.0290***	0.0441***	0.0242***
	(0.0000)	(0.0000)	(0.0000)
Stage 3	-0.0324***	-0.0277***	-0.0392***
	(0.0000)	(0.0000)	(0.0000)
Stage 4	-0.0431***	-0.0423***	-0.0479***
	(0.0000)	(0.0000)	(0.0000)
Education *Stage 3	-0.0111***	-0.0225***	-0.0030***
	(0.0000)	(0.0000)	(0.0000)
Education *Stage 4	-0.0174***	-0.0358***	0.0158***
	(0.0000)	(0.0000)	(0.0000)
Intercept	0.1496	0.1237	0.1689
N	20000	10000	10000

Standard errors in parentheses. Two-tailed test.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The results in Table7are in line with the findings by Santaeulalia-Llopis (2016). The HIV educational gradient in the myopic stage of the epidemic is significantly positive, both in the total population as well as the gender dis aggregation. The rest of the coefficients are significantly negative, meaning that in later stages of the epidemic, additional education represents a reduction of the probability of infection. The HIV education gradient by stage of the epidemic is summarized in Table8 and graphed in Figure1 and Figure2.

Table 8: HIV-education gradient by stage of the epidemic and by gender(%)

HIV-Education Gradient	**Total**	**Males**	**Females**
Myopic stage(γ_2)	2.9	4.41	2.42
Maturity stage$(\gamma_2 + \alpha_3 + \gamma 3)$	-1.45	-0.60	-1.80
ARTs stage$(\gamma_2 + \alpha_4 + \gamma 4)$	-3.16	-3.39	-0.79

The results in Table8 can be interpreted as follows. In the *myopic stage* people who finish secondary education or have more years of education, increase their probability of getting HIV infected by 2.9%. However on later stages of the epidemic, namely the *myopic stage* and the *maturity stage*, increasing years of education beyond secondary education reduces the chances of HIV infection by 1.45% and 3.16% respectively. When looking at the gender dis-aggregation, we can see that education beyond secondary increases the probability of HIV infection in the *myopic stage* by 4.41% but in the *ARTs stage*, additional education, reduces the risk oh HIV infection by 3.39%. The same occurs in the case of females, but the reduction is however the probability of infection does not reduce as much as in the two other cases.

Figure1 shows that the educational gradient is concave up, and starts to have the U shape described by Santaeulalia-Llopis (2016). Nevertheless, in Malawi the HIV-education gradient starts positive in the early stages, but turns significantly negative in the last two stages. This suggest that it is wise to encourage education in the latest stages of the epidemic, as this will reduce the probabilities of infection in the future.

The HIV education gradient for females has a clear U shape. The gradient start at 2.42 % during the *myopic stage* but then turns negative to rise again after reaching its lower value in the during the *maturity stage* . This means that educated women in the latest stages should be considered as a vulnerable target group. When looking at the gradient of men, we see that it has a clear downward sloping shape. More precisely, additional years of education for men will reduce their chances of HIV infection. Then, targeted intervention should include all those individuals who possess less education. We can also see that the HIV gradient for females is larger in magnitude than that of men. In neither of the cases the gradient turns to positive during the *ARTs stage* . This a big difference when comparing with the results obtained for Sub-Saharan Africa. [11]

Figure 1: HIV-education gradient total

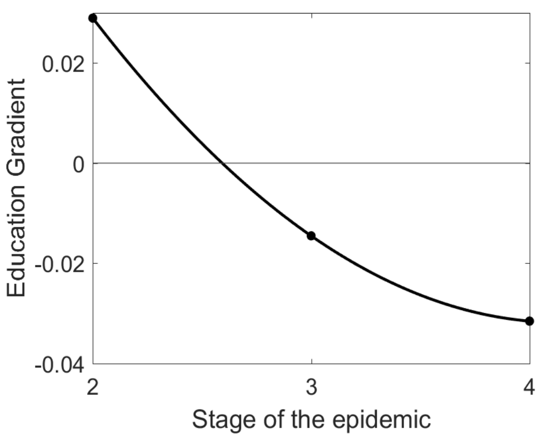
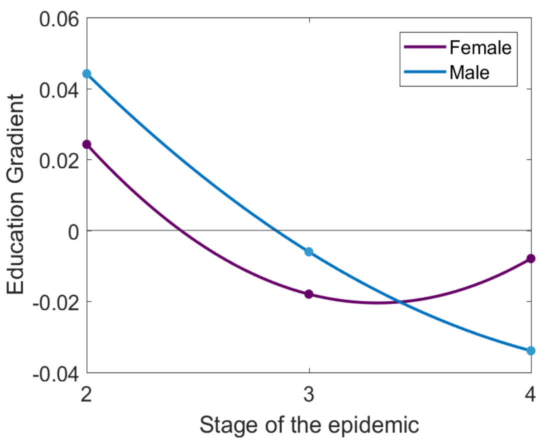


Figure 2: HIV-education gradient by gender



In order to understand the differences between educated risky sex buyers and risky

[11]See Figure5 to Figure6 in the AppendixA

sex producers, it is possible to conduct a simple sensitivity analysis on the estimation results. One way to increase the overall prevalence in the country is to increase labor income given to all individuals, in particular, any individual regardless if she/he is infected or not, will increase its chosen level of extramarital risky sex no matter what because of the income effect, as a result, overall prevalence among the population will increase. A second channel that increases prevalence among the educated buyers or producers is a reduction of the "comprehension parameter" ρ_e, since educated agents are less aware then they consume more risky sex. However these channels don't tell us anything about the prevalence by gender. On the other hand, there is a set endogenous outcomes generated by the sex market behavior and the amount of risky sex demanded by educated/uneducated sex buyers. In principle, if the price of risky sex increases, the total demand will decrease, how ever total demand might is still covered be covered in the same proportion by educated/uneducated risky sex producers. If demand for risky sex among uneducated sex buyers increases, it might be the case that prevalence among educated women increases since they also have to cover the increasing demand. In other words the market protects educated sex buyers, since in no way they will be forced to increase their sex consumption, but if uneducated buyers increase demand, both uneducated and educated producers must intervene to clear the sex market.

7 Conclusions

In this paper I propose a general equilibrium heterogeneous agent model intended to help understand the dynamic relation between education and HIV diffusion. The model features two types of agents that differ primarily in their level of education, asset holdings and HIV status. The model identifies four stages of the HIV-epidemic: 1.A *pre-epidemic stage*, where there is no risk of infection. 2. A *myopic stage* where individuals do not realize the hazard of HIV infection and don't take it into account for consumption decisions. 3.A *maturity stage* where individuals acknowledge the risk of infection and immediately adjust their expectations and consumption decisions. It is because of this reason that the equilibrium amount of extramarital risky sex in the maturity of the epidemic is less than that the *myopic stage*. Finally the *ARTs stage* introduces antiretroviral treatment that might cause the prevalence among educated people to increase. This can be explained by the fact that since educated individuals believe they have a lower risk of infection, then they start to increase their levels of extramarital risky sex again and this has overall positive effect on prevalences.

The paper proposes an innovative algorithm that links all four epidemic stages together in order to have a dynamic representation of the HIV epidemic. The model is then calibrated for Malawi using mostly DHS data. Once the model is solved, I use the data generated by the model to calculate the HIV education gradient for Malawi and its dis aggregation by gender.

The gradient has an evident U shape, however, after the *myopic stage* the gradient becomes significantly negative, meaning that it is recommendable to increase the level of education of the population in order to reduce the risk of HIV infection. Moreover when looking at the gender dis-aggregation we note that education indeed reduces the probability of infection both among men and women. But for the case of women after the *maturity of the epidemic*, the education gradient starts to get closer to zero. This suggests that increasing the level of education of the female population in Malawi will not reduce HIV prevalence as in the *maturity of the epidemic*. Quantitatively speaking, during the *myopic stage*of the epidemic additional education increases the probability of infection by 2.9% among the total population and among males and females 4.41% and 2.42 % respectively. However during the *ARTs stage* education actually reduces the risk of infection by 3.16% among the total population and among men and women 3.39% and 0.79 % respectively.

References

Alsan, M. and Cutler, D. M. (2013). Girls education and hiv risk: Evidence from uganda. *Journal of Health Economics*, 32(5):863–872.

Avert (2018). Hiv and aids in malawi. www.avert.org/professionals/hiv-around-world/sub-saharan-africa/malawi.

Beegle, K. and de Walque, D. (2009). Demographic and socioeconomic patterns of hiv/aids prevalence in africa. Policy Research Working Paper Series 5076, The World Bank.

Bethencourt, C. and Rios-Rull, J.-V. (2009). On the living arrangements of elderly widows. *International Economic Review*, 50(3):773–801.

Bledsoe, C. (1990). Transformations in sub-saharan african marriage annd fertility. *The annals of the American Academy of Political and Social Science*, 510:115–125.

Brent, R. (2006). Does female education prevent the spread of hiv-aids in sub-saharan africa? *Applied Economics*, 38(5):491–503.

Case, A. and Paxson, C. (2013). Hiv risk and adolescent behaviors in africa. Working papers, Princeton University, Woodrow Wilson School of Public and International Affairs, Center for Health and Wellbeing.

de Walque, D. (2007). How does the impact of an hiv/aids information campaign vary with educational attainment? evidence from rural uganda. *Journal of Development Economics*, 84(2):686–714.

Fortson, J. (2008). The gradient in sub-saharan africa: Socioeconomic status and hiv/aids. *Demography*, 45(2):303–322.

Greenwood, J., Kircher, P., Santos, C., and Tertilt, M. (2017). An equilibrium model of the african hiv/aids epidemic. RCER Working Papers 601, University of Rochester - Center for Economic Research (RCER).

Kremer, M. (1996). Integrating behavioral choice into epidemiological models of aids. *The Quarterly Journal of Economics*, 111(2):549–573.

Manuelli, R. (2015). Aids, human capital and development. 2015 Meeting Papers 1193, Society for Economic Dynamics.

Mishra, V., Assche, S., Greener, R., Vaessen, M., Hong, R., Ghys, P., Boerma, J., Van Assche, A., Khan, S., and Rutstein, S. (2007). Hiv infection does not disproportionately affect the poorer in sub-saharan africa. *Journal of Development Economics*, 84(2):686–714.

Oster, E. (2012). Routes of infection: Exports and hiv incidence in sub-saharan africa. *Journal of the European Economic Association*, 10(5):1025–1058.

Preston, S. H. (1996). American Longevity: Past, Present, and Future. Center for Policy Research Policy Briefs 7, Center for Policy Research, Maxwell School, Syracuse University.

Santaeulalia-Llopis, R. (2016). Education, hiv status and risky sexual behavior: How much does the stage of the hiv epidemic matter ? Economics Working Papers ECO2016/09, European University Institute.

UNAIDS (2003). Hiv /aids at glance. World Bank Health-Nutrition-Population. www.unaids.org.

Yao, Y. (2016). Fertility and hiv risk in africa. Working Paper Series 5342, Victoria University of Wellington, School of Economics and Finance.

A Appendix

A.1 Stationary equilibrium computation

The follow algorithm explains the steps to find the stationary equilibrium of each of the stages of the epidemic. The enviroment is characterized by the stages $i = (1, 2, 3, 4)$ explained in Section5.

Initialization

Step 1**:**

Guess prices p, r.

1.1. We are in Stage i.

1.2. Use the first order conditions of the agents in Stage i to solve the problems for sex buyers and sex producers.

1.3. Use the decision rules and the associated transition function $\mathcal{Q}((a, e, g, h)(\mathcal{A}, \mathcal{E}, \mathcal{G}, \mathcal{H}))$ to find an invariant distribution $\Phi(\mathcal{A}, \mathcal{E}, \mathcal{G}, \mathcal{H})$.

1.4. Compute:

$$Ea_g(p, r) = \int_{a,e,g,h} a(a, e, g, h; \Phi) d\Phi \ \forall g \in \mathcal{G}$$
$$Ex_g(p, r) = \int_{a,e,g,h} x(a, e, g, h; \Phi) d\Phi \ \forall g \in \mathcal{G}$$

1.5. Compute:

$$d^a(r, p) = Ea_g(p, r) - Ea_{-g}(p, r)$$
$$d^x(r, p) = Ex_g(p, r) - Ex_{-g}(p, r)$$

Step 2**:**

2. If both $d^a(r, p), d^x(r, p)$ are substantially different than zero, update r, p and go back to Step 1.1. Otherwise stop.

A.2 Figures

A.2.1 Education gradient Sub Saharan Africa

Figure 3: Gradient Whole population

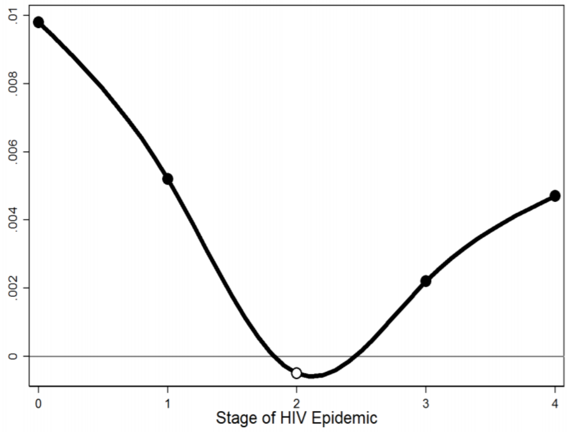
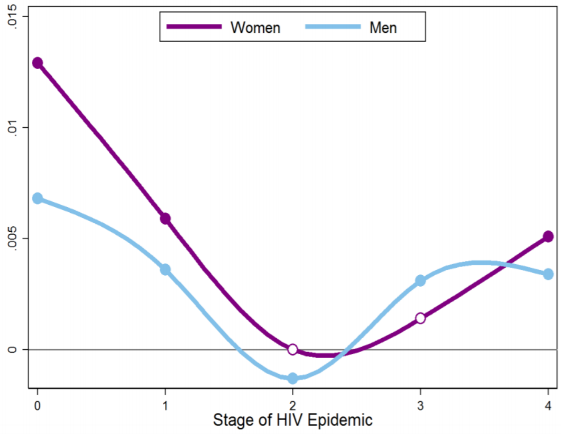
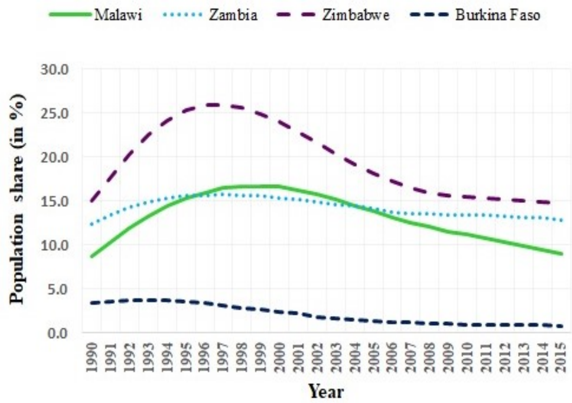


Figure 4: Gradient Gender disaggregation



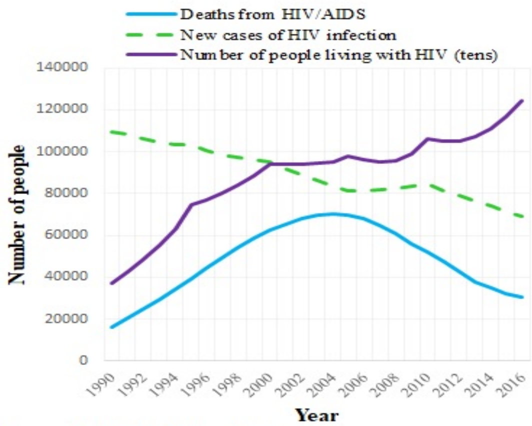
A.2.2 Malawian HIV/AIDS epidemic

Figure 5: Share of people living with HIV



Source: World bank, World development indicators

Figure 6: Malawi, people living with HIV, new infections and deaths due to HIV



Source: IHME, Global burden disease

Figure 7: Malawi, AIDS-related deaths among adults by gender

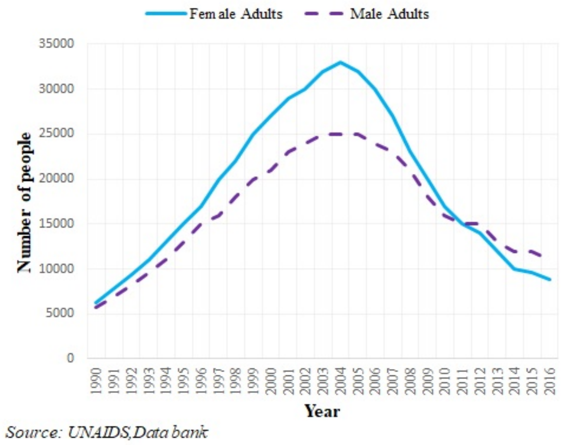
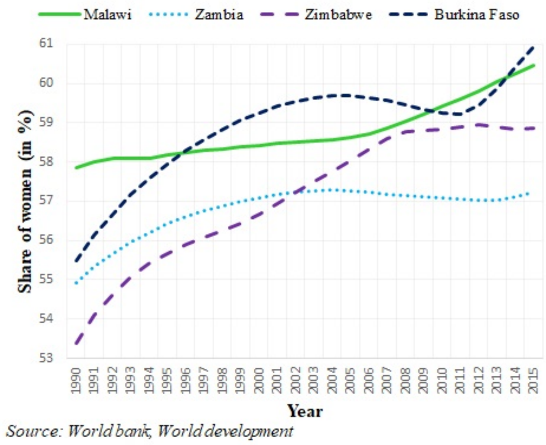


Figure 8: Share of Women among the population living with HIV



A.2.3 Pre-epidemic stage

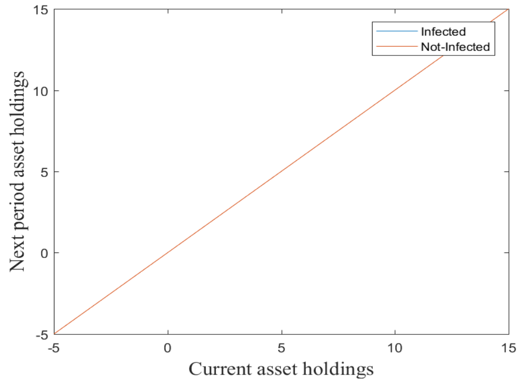
A.2.4 Miopic stage of the epidemic

A.2.5 Maturity of the epidemic

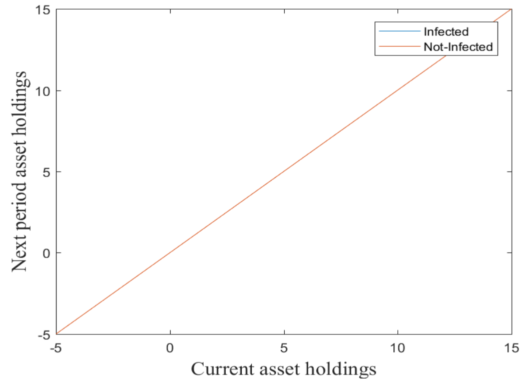
A.2.6 Antiretrovial treatment (ARTs)

Figure 9: Pre-epidemic stage

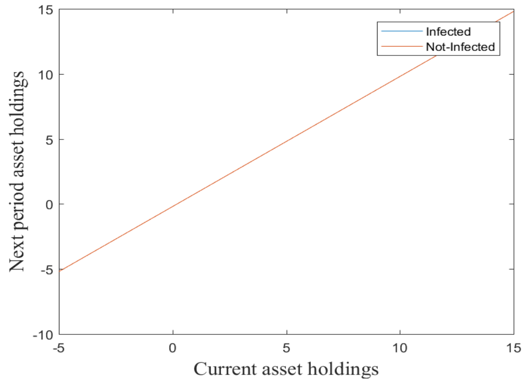
(a) PF asset holdings (Edu-Buyers)



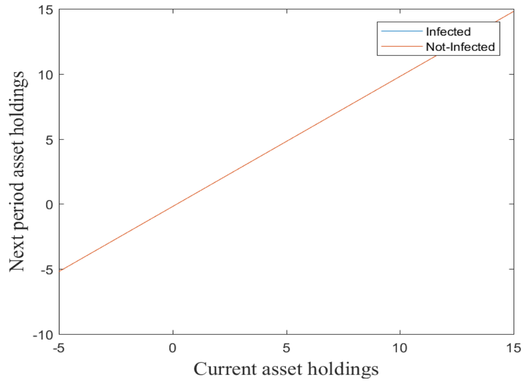
(b) PF asset holdings (Unedu-Buyers)



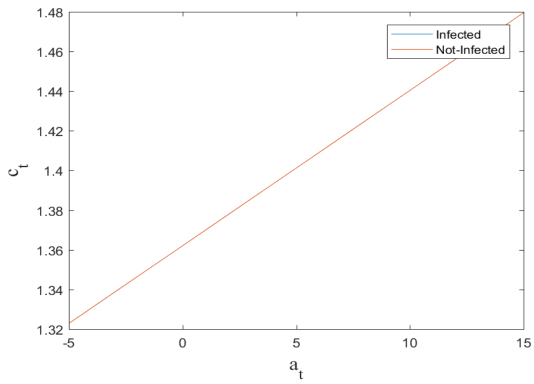
(c) PF asset holdings (Edu-Produ)



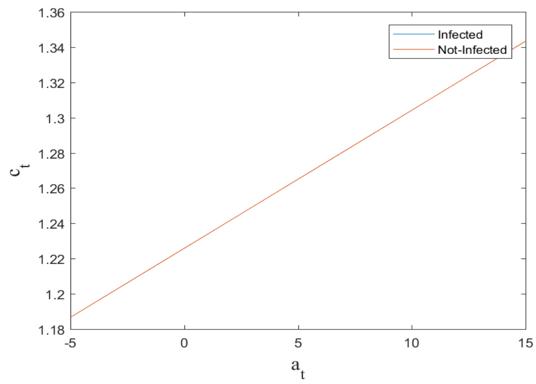
(d) PF asset holdings (Unedu-Produ)



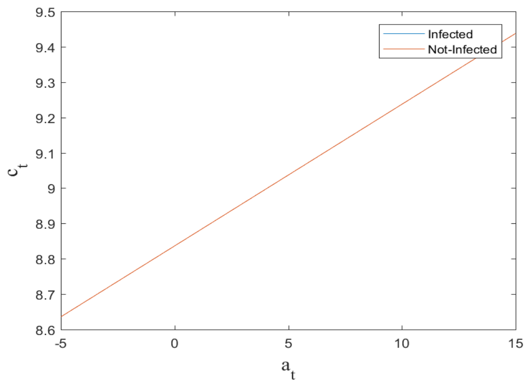
(e) PF consumption (Edu-Buyers)



(f) PF consumption (Unedu-Buyers)



(g) PF consumption (Edu-Produ)



(h) PF consumption (Unedu-Produ)

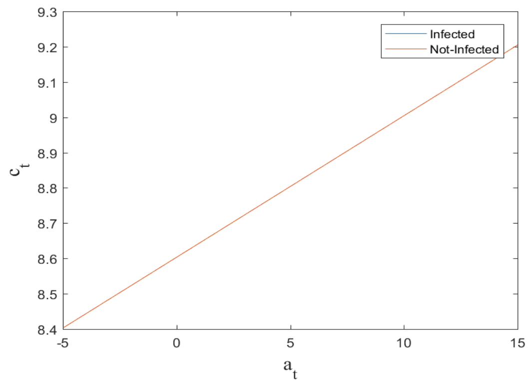
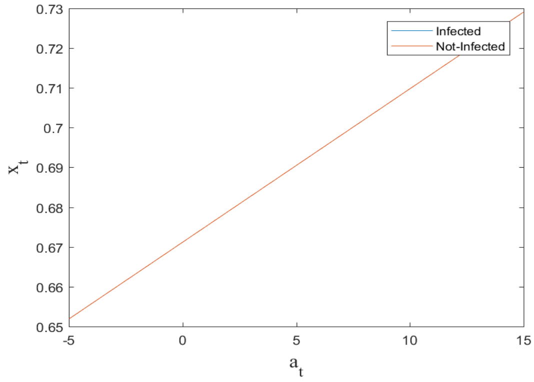
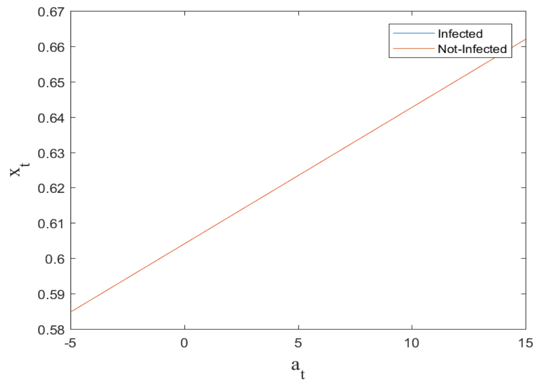


Figure 10: Pre-epidemic stage, continuation

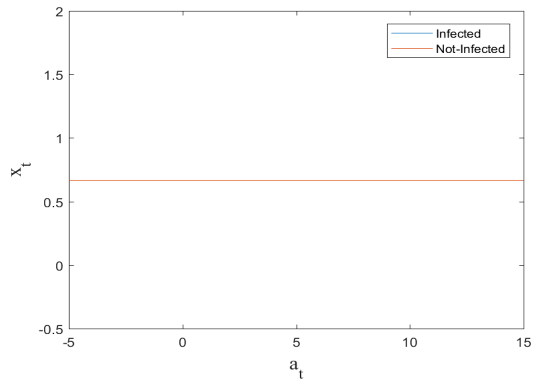
(a) PF sex consumption (Edu-Buyers)



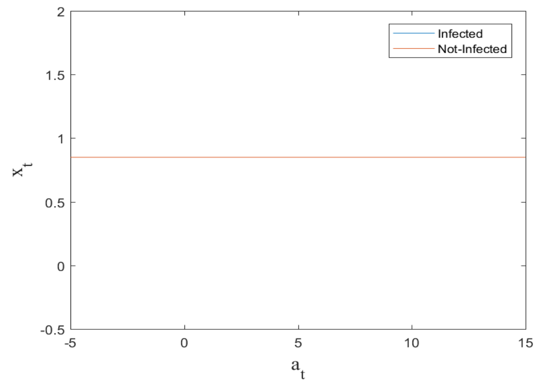
(b) PF sex consumption (Unedu-Buyers)



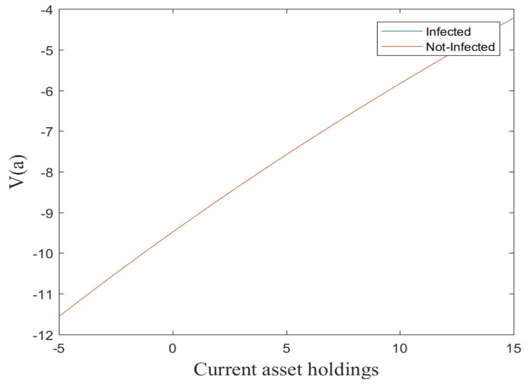
(c) PF sex production (Edu-Produ)



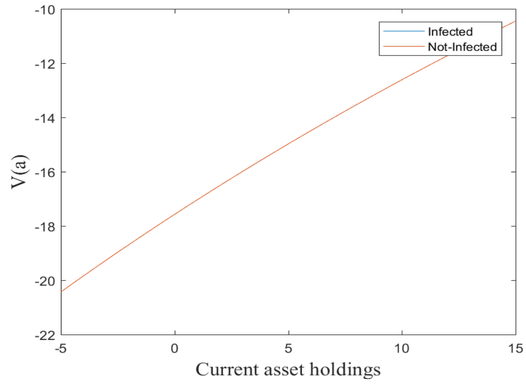
(d) PF sex production (Unedu-Produ)



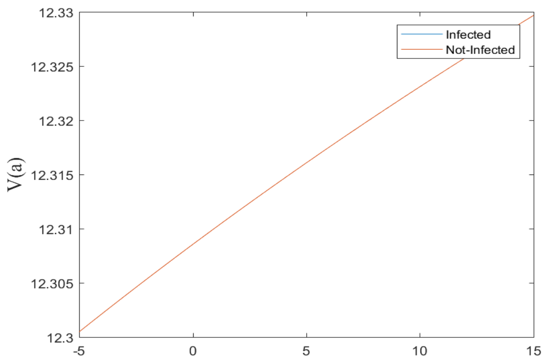
(e) Value Function (Edu-Buyers)



(f) Value Function (Unedu-Buyers)



(g) Value Function (Edu-Produ)



(h) Value Function (Unedu-Produ)

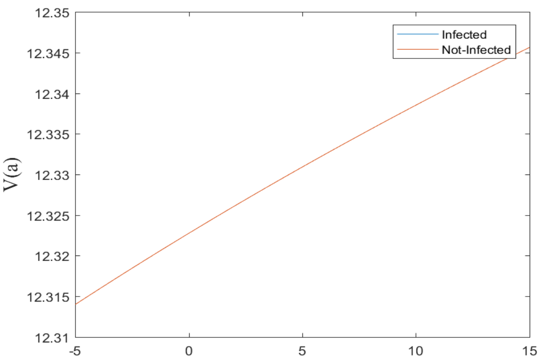
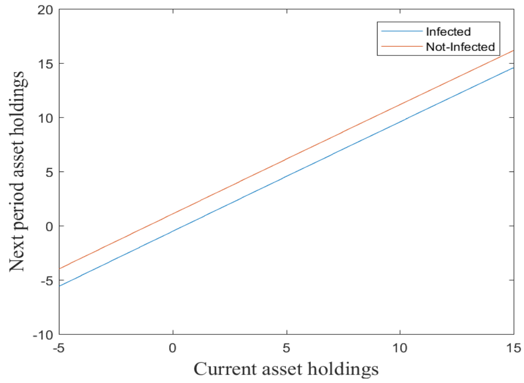
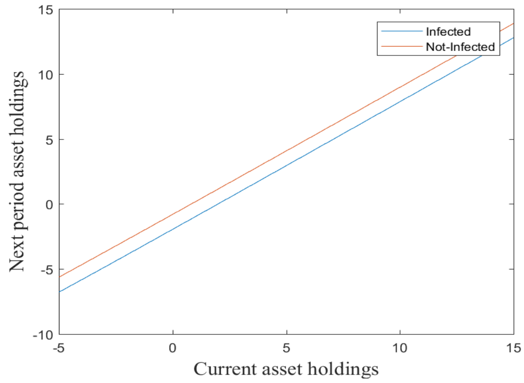


Figure 11: Myopic onset of the epidemic

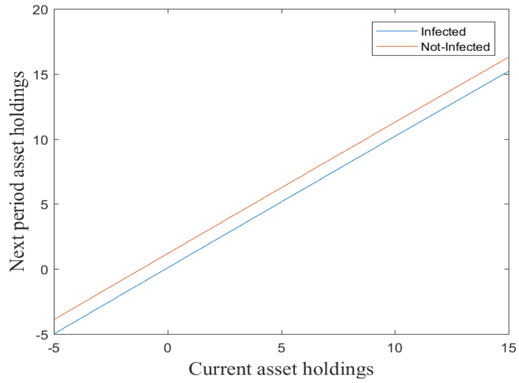
(a) PF asset holdings (Edu-Buyers)



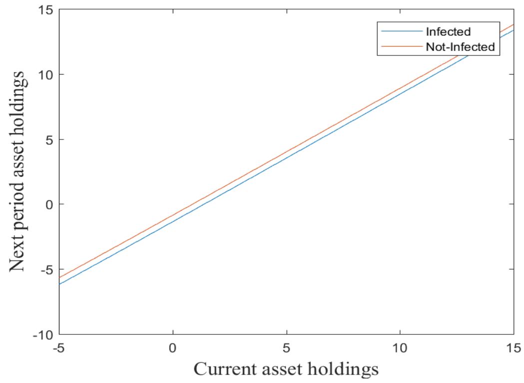
(b) PF asset holdings (Unedu-Buyers)



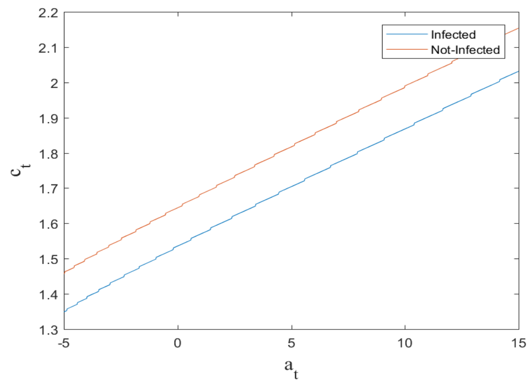
(c) PF asset holdings (Edu-Produ)



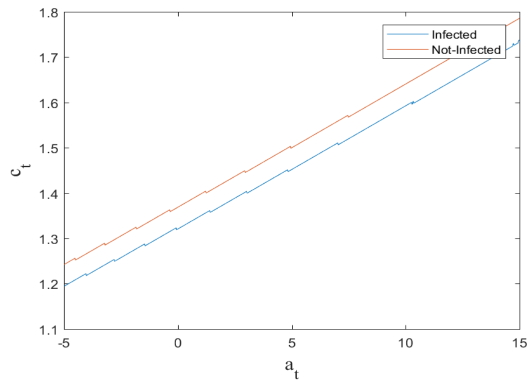
(d) PF asset holdings (Unedu-Produ)



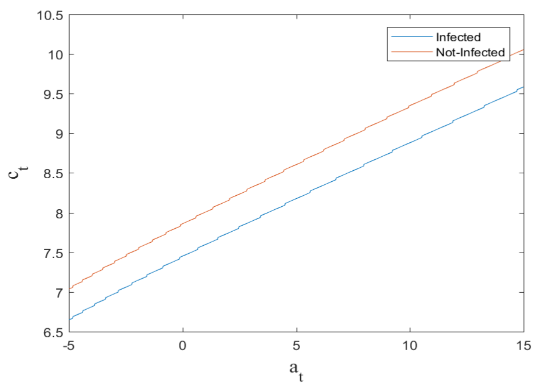
(e) PF consumption (Edu-Buyers)



(f) PF consumption (Unedu-Buyers)



(g) PF consumption (Edu-Produ)



(h) PF consumption (Unedu-Produ)

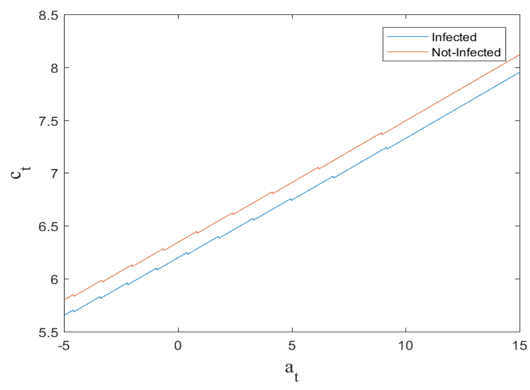
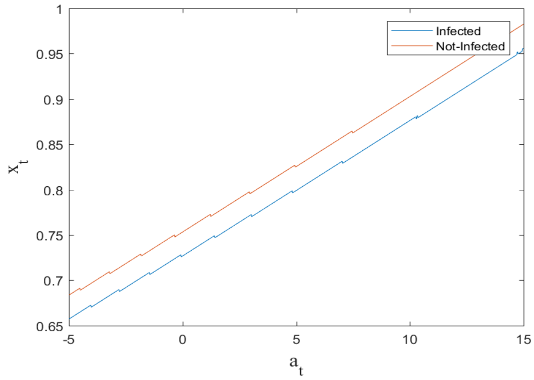
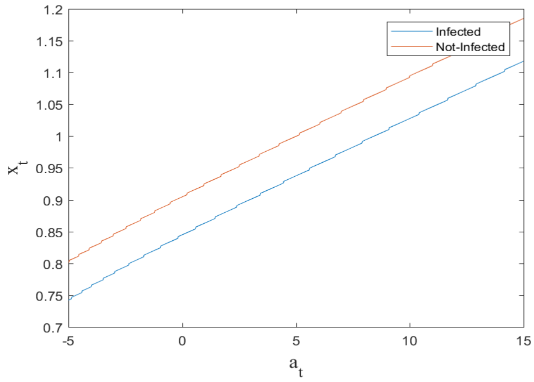


Figure 12: Myopic onset of the epidemic, continuation

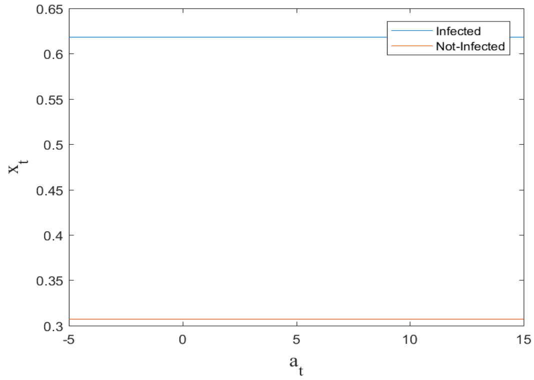
(a) PF sex consumption (Edu-Buyers)



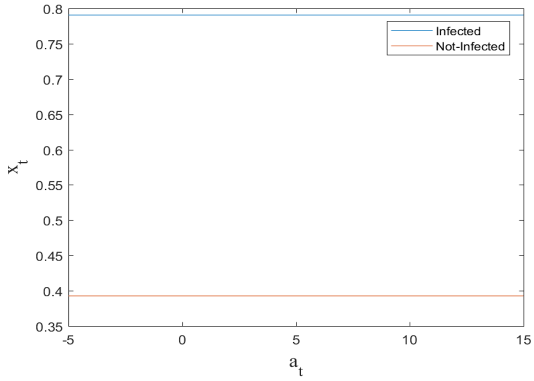
(b) PF sex consumption (Unedu-Buyers)



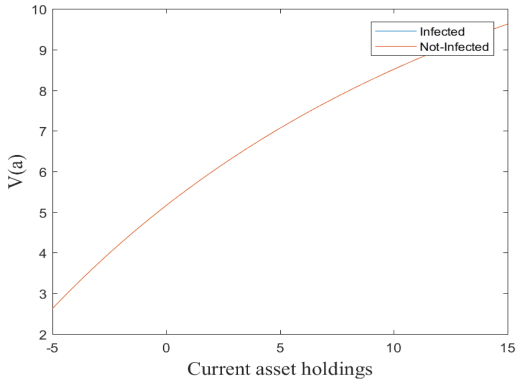
(c) PF sex production (Edu-Produ)



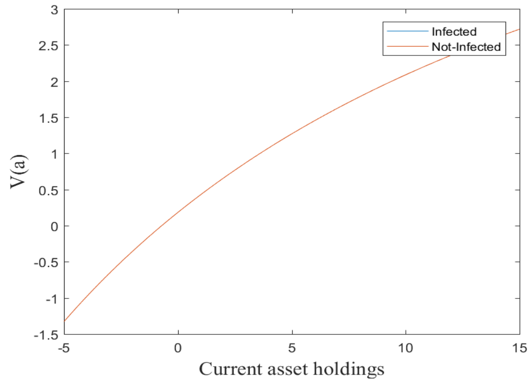
(d) PF sex production (Unedu-Produ)



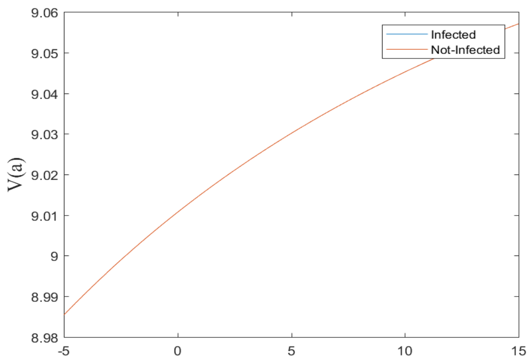
(e) Value Function (Edu-Buyers)



(f) Value Function (Unedu-Buyers)



(g) Value Function (Edu-Produ)



(h) Value Function (Unedu-Produ)

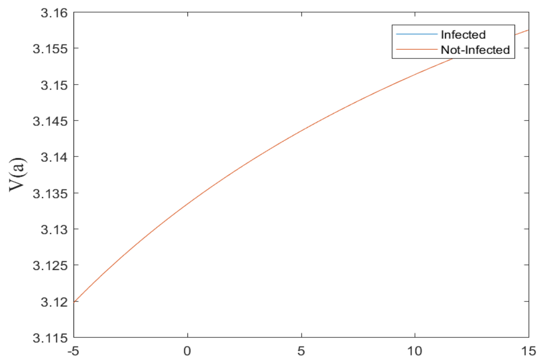
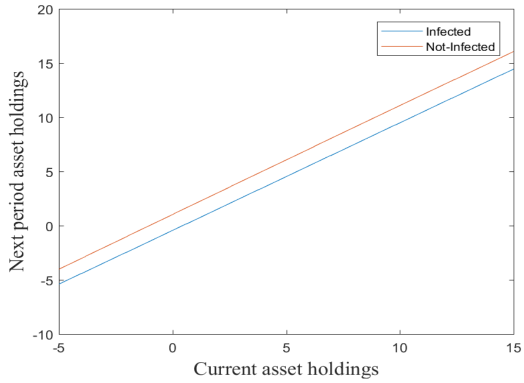
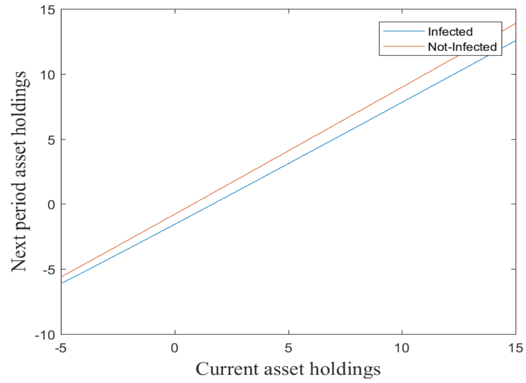


Figure 13: Maturity and diffusion

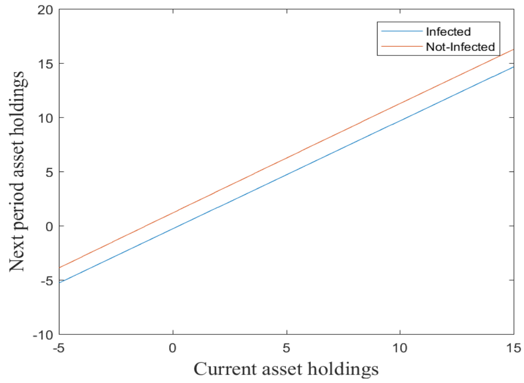
(a) PF asset holdings (Edu-Buyers)



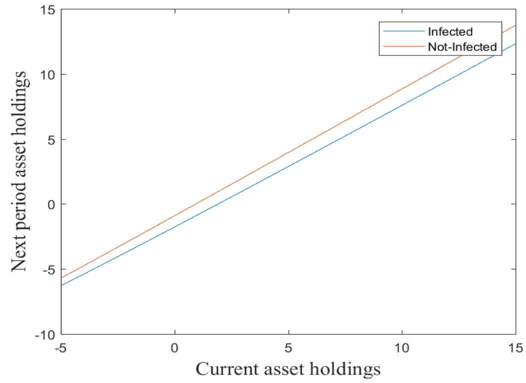
(b) PF asset holdings (Unedu-Buyers)



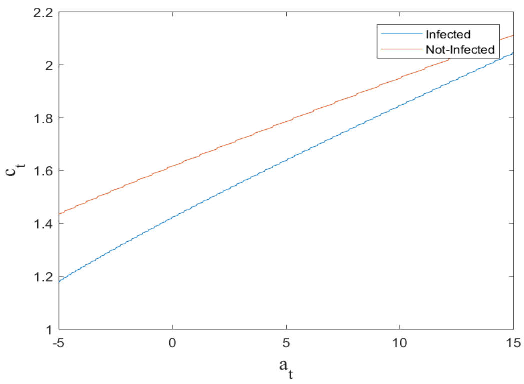
(c) PF asset holdings (Edu-Produ)



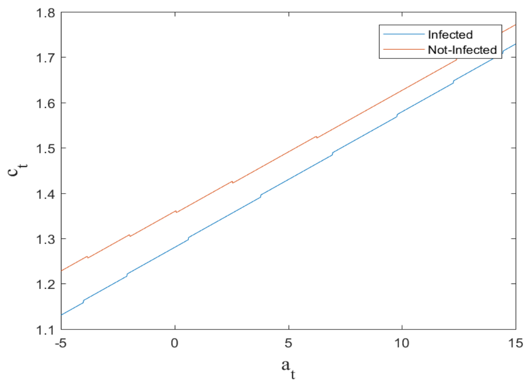
(d) PF asset holdings (Unedu-Produ)



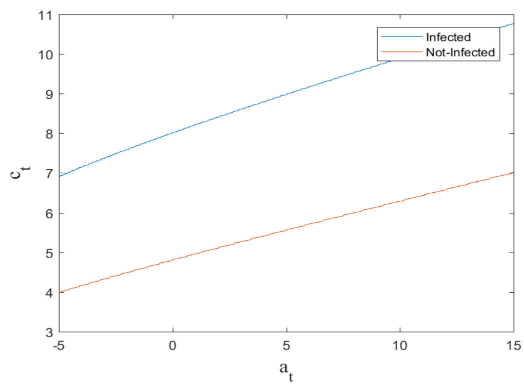
(e) PF consumption (Edu-Buyers)



(f) PF consumption (Unedu-Buyers)



(g) PF consumption (Edu-Produ)



(h) PF consumption (Unedu-Produ)

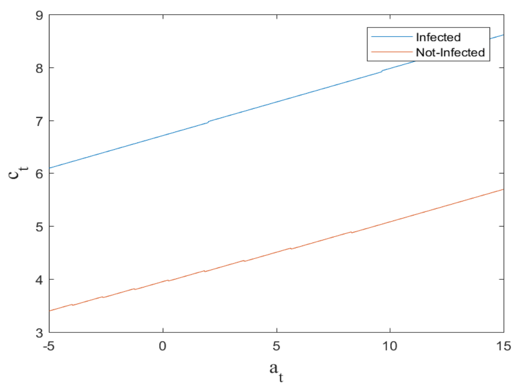
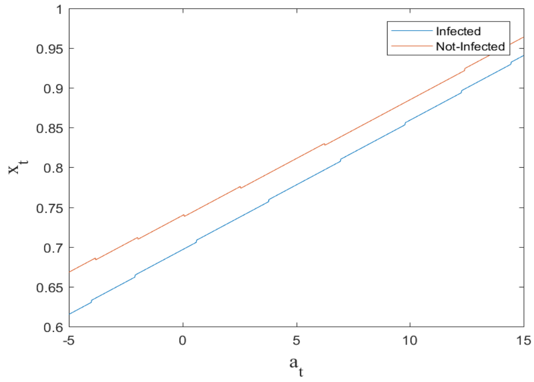
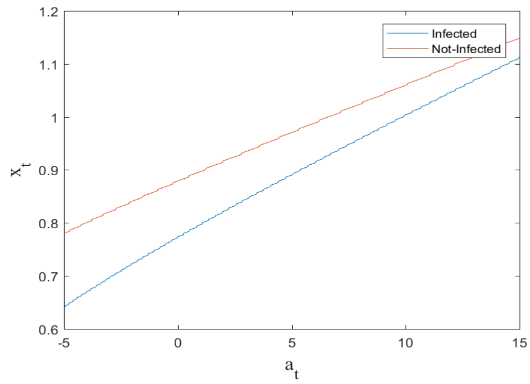


Figure 14: Maturity and diffusion, continuation

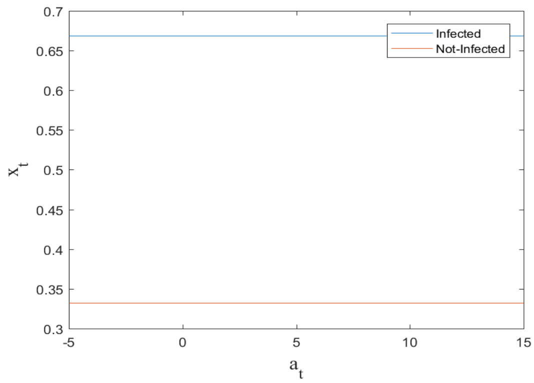
(a) PF sex consumption (Edu-Buyers)



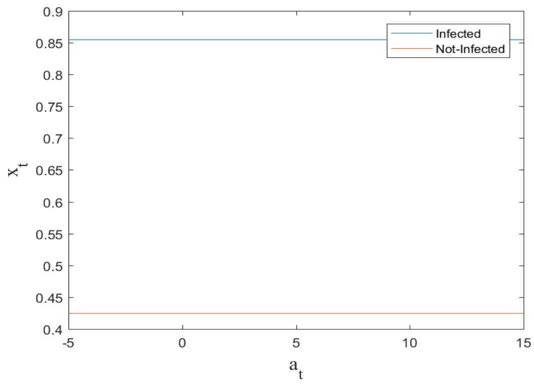
(b) PF sex consumption (Unedu-Buyers)



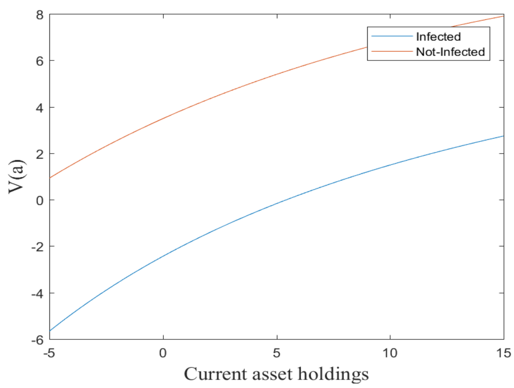
(c) PF sex production (Edu-Produ)



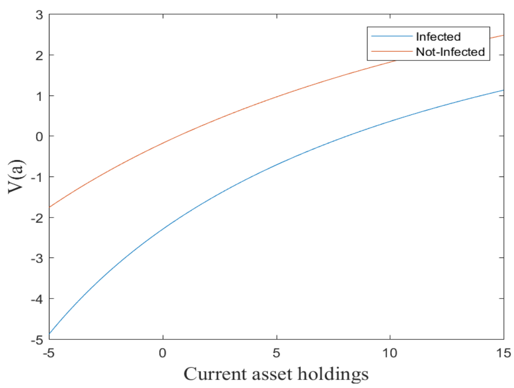
(d) PF sex production (Unedu-Produ)



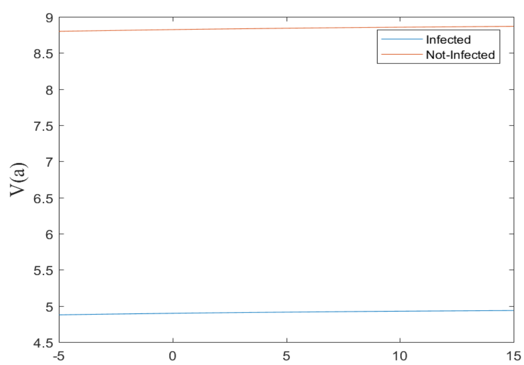
(e) Value Function (Edu-Buyers)



(f) Value Function (Unedu-Buyers)



(g) Value Function (Edu-Produ)



(h) Value Function (Unedu-Produ)

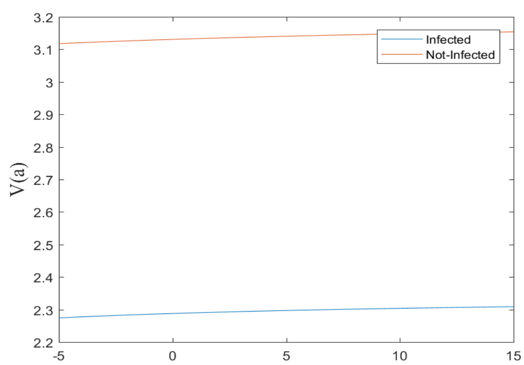
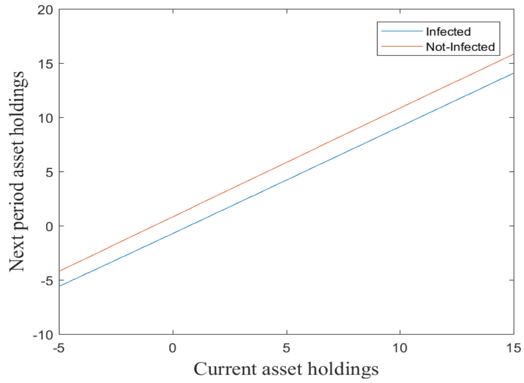
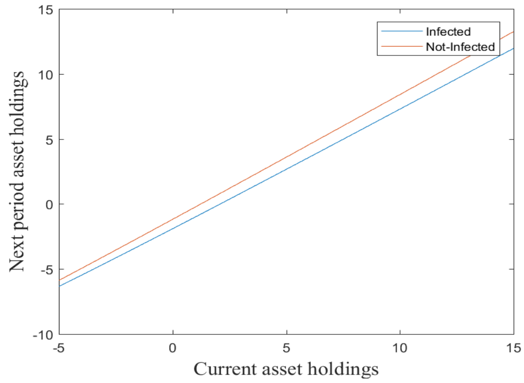


Figure 15: Antiretroviral treatment (ARTs)

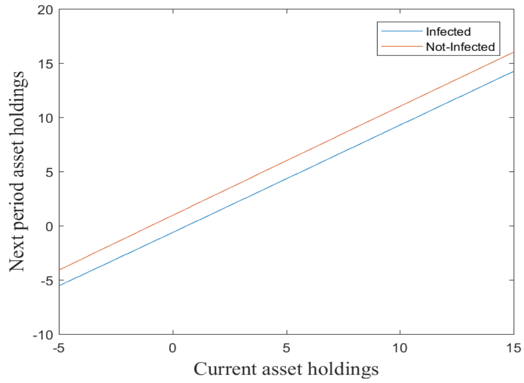
(a) PF asset holdings (Edu-Buyers)



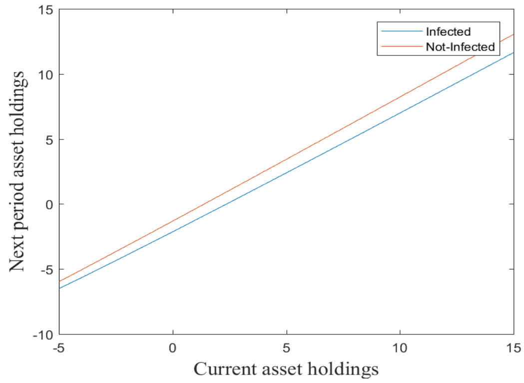
(b) PF asset holdings (Unedu-Buyers)



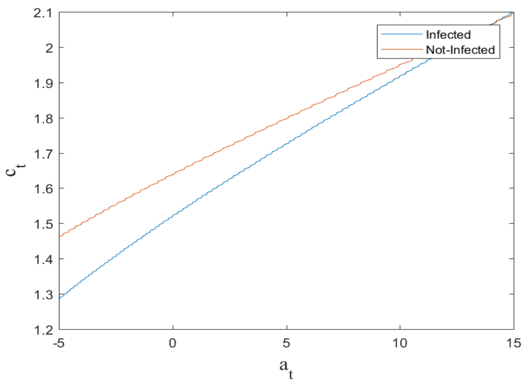
(c) PF asset holdings (Edu-Produ)



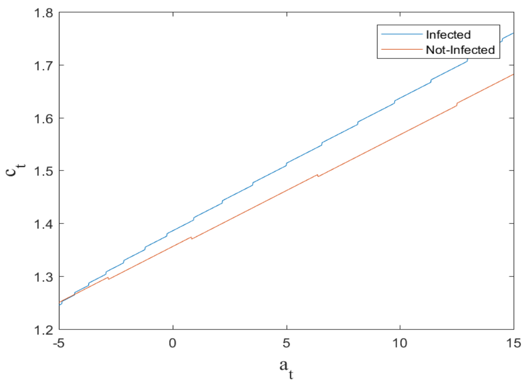
(d) PF asset holdings (Unedu-Produ)



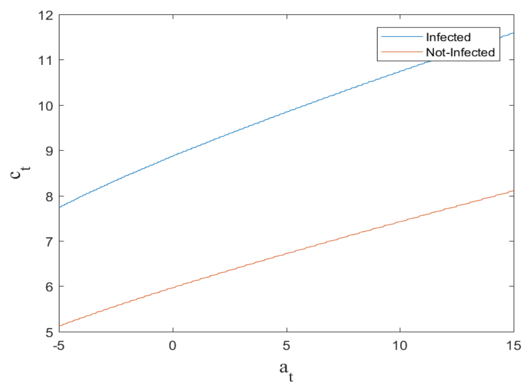
(e) PF consumption (Edu-Buyers)



(f) PF consumption (Unedu-Buyers)



(g) PF consumption (Edu-Produ)



(h) PF consumption (Unedu-Produ)

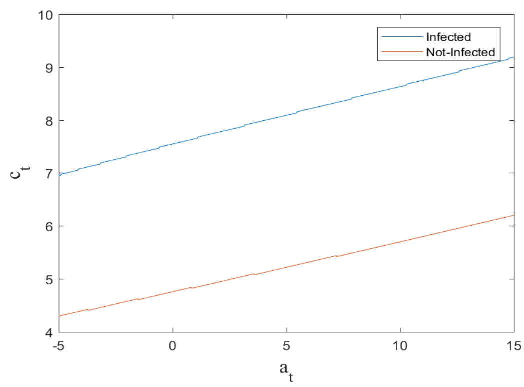
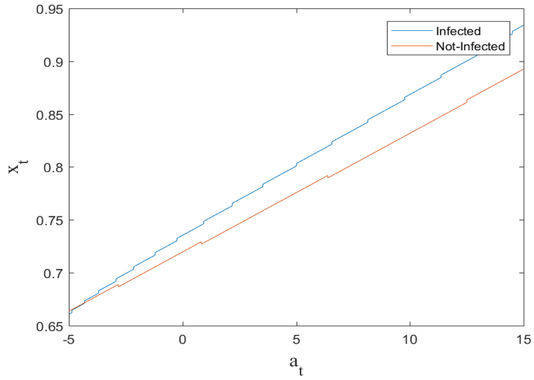
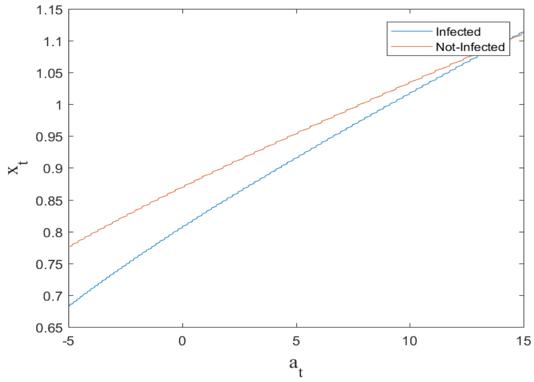


Figure 16: Antiretroviral treatment (ARTs), continuation

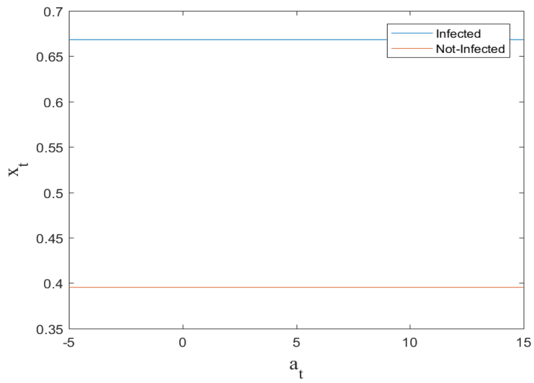
(a) PF sex consumption (Edu-Buyers)



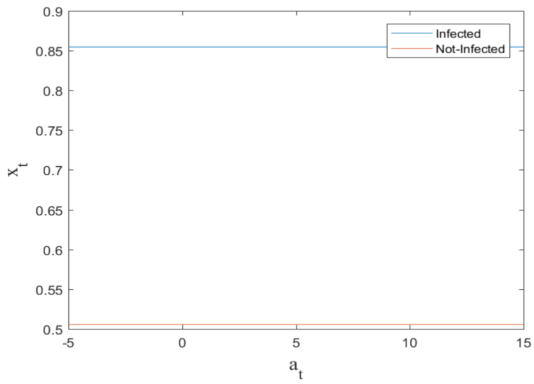
(b) PF sex consumption (Unedu-Buyers)



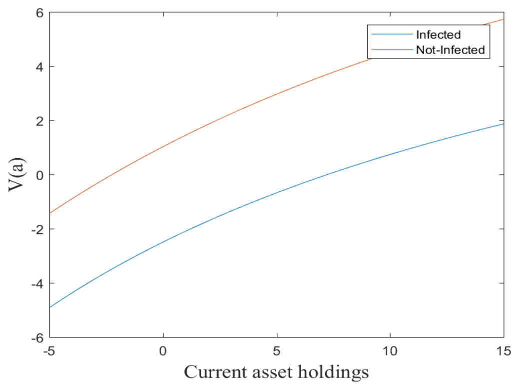
(c) PF sex production (Edu-Produ)



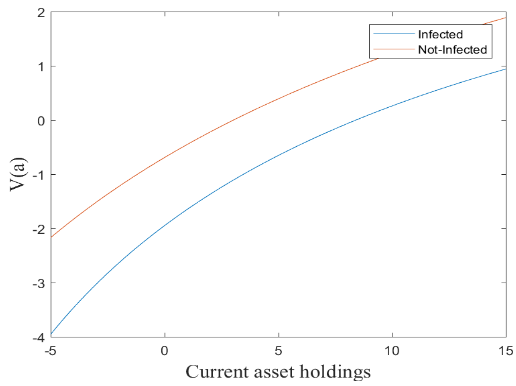
(d) PF sex production (Unedu-Produ)



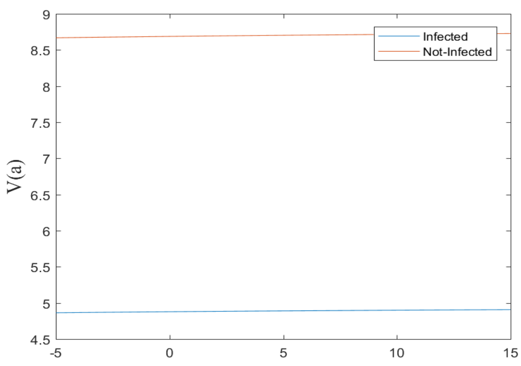
(e) Value Function (Edu-Buyers)



(f) Value Function (Unedu-Buyers)



(g) Value Function (Edu-Produ)



(h) Value Function (Unedu-Produ)