Department of Economics- FEA/USP

Marriage in the time of the HIV/AIDS epidemic

LORENA HAKAK

PAULA PEREDA

WORKING PAPER SERIES Nº 2021-07

DEPARTMENT OF ECONOMICS, FEA-USP

WORKING PAPER Nº 2021-07

Marriage in the time of the HIV/AIDS epidemic

Lorena Hakak (lhakak@usp.br)

Paula Pereda (pereda@usp.br)

Abstract:

HIV/AIDS was the main cause of death among young adults in the 1990s. The sexual freedom from the rise of contraceptives and women's empowerment affected individuals' preferences for dating, marriage and fertility. In this paper, we investigate whether the HIV/AIDS epidemic from the 1980s onward, a non-monetary shock, affected people's preferences for marriage. We set out a simple two period non-cooperative model and demonstrate that under a growing likelihood of meeting someone infected by AIDS, the expected payoff of people who got married in the first period increased. Then, using unique Brazilian data, we estimate the marriage gain in two different time spans, 1984 to 1991 and 1999 to 2010, respectively before and after the free distribution of antiretrovirals by the Brazilian government. Our findings corroborate that HIV/AIDS influenced marriages in the past, working as a response against the risk of being infected and increasing the marital surplus, especially for women. We also find evidence that the impact of HIV/AIDS on marriage disappeared once antiretroviral drugs were universally distributed, which significantly reduced the mortality risk of the disease, and therefore the expected costs of the disease. The perceived vulnerability to the consequences of HIV/AIDS decreased, resulting in a change in sexual behavior, thus reducing the value of monogamous relationships compared to the period without antiretroviral drugs.

Keywords: marriage Market; family economics; HIV/AIDS; antiretroviral treatment.

JEL Codes: D1; J12; I1

Marriage in the time of the HIV/AIDS epidemic

Lorena Hakak * Paula Pereda†

October 21, 2022

Abstract

We investigate whether the non-monetary shock of the HIV/AIDS epidemic affected people's preferences for marriage. We set out a two-period non-cooperative model and demonstrate that under the growing likelihood of meeting someone infected by AIDS, the expected payoff of people who got married in the first period increased. Then, using unique data for Brazil, we estimate the static and dynamic marital surplus before and after the free distribution of antiretrovirals by the Brazilian government. Our findings corroborate that the epidemic increased the value of marriage in the past, especially for highly educated couples and for women. We find evidence that the impact of HIV on marriage disappeared once the government freely distributed antiretroviral drugs, which has changed the risk perception of the disease.

JEL codes: D1, J12, I1.

Keywords: marriage market, family economics, HIV/AIDS, antiretroviral treatment, sexual behavior.

*Federal University of ABC - UFABC and GeFam. Contact: lorena.hakak@ufabc.edu.br

†University of Sao Paulo - FEA/USP. The authors acknowledge the comments from Vladimir Ponczek, Emerson Marçal, Rebecca Myerson, Cezar Santos, Aloysius Siow, Edson Severnini, Sergio Firpo, Anastasia Terskaya, Shelly Lundberg, Jeanne Lafortune and Rodrigo Soares and seminar participants at 3rd Brazilian Meeting on Family and Gender Economics (GeFam/2022), Western Economic Association International (WEAI/2019), the 2019 Annual Congress of the European Economic Association (EEA/2019), and the Regional and Urban Economics Lab (NEREUS) academic seminars at University of Sao Paulo. We are particularly thankful to Vladimir Ponczek for his comments on the instrument we use. Prof. Pereda also thanks the research support from Fipe. Remaining errors are ours.

1 Introduction

Acquired immunodeficiency syndrome (AIDS)[1] was the main global cause of death in young adults at the end of the past century[2]. The evolution of the mortality risk by HIV/AIDS over time might have affected the demand for stable relationships. Single people may change their sexual behavior by looking for more stable relationships, and ultimately a life partner, making decisions on marriage sooner. In this case, the utility of remaining single decreases whereas marital surplus increases. Incentives for marriage might change, causing the marriage market to create a new shadow price to guide individuals to marriage so as to maximize their expected well-being (Becker 1991). As far as we know, there is no empirical research examining whether this sexually transmitted disease (health shock) influenced the value of marriage. In this sense, this paper aims to fill in this gap.

We investigate whether the HIV/AIDS epidemic changed preferences for marriage. First, we present a theoretical motivation based on a two-period non-cooperative model. We analyze the channel of sex behavior, under the increase in likelihood of dating someone infected by AIDS, with the marital surplus. We demonstrate that under the growing likelihood of meeting someone infected by AIDS, the expected payoff of people who got married in the first period increases. Then, we propose an empirical model to study individuals' marriage behavior and its relationship with HIV/AIDS infections using Brazilian data. The country is an interesting setting for this analysis for three reasons: (i) it was hardly affected by the AIDS epidemic[3], (ii) the number of marriages and cohabitation changed direction in 1984, as the number of reported cases of AIDS had increased, and (iii) after 1996, Brazilian government was the first to provide antiretroviral drugs (ARVs) free of charge for infected people.

The number of marriages and cohabitation, which had been falling in the beginning of the 80's, changed direction and had strongly increased until 1986 (see Panels (a) and (c) of Figure 1). This turning point coincided with the increase in the number of reported cases of AIDS in the country. From 1987 to 1991, the total number of marriages and cohabitation remained practically stable, while formal unions fell. There seems to have been a kind of substitution between formal and informal unions, the latter very common in the lower classes.

[1] AIDS is the disease caused by infection by the Human Immunodeficiency Virus (HIV).

[2] According to the Global Burden of Diseases data for individuals of both sexes aged between 15-49 years from 1990 to 2000 (Murray et al. 2018). Available at `https://vizhub.healthdata.org/gbd-compare/`.

[3] The number of men (women) reported cases of AIDS were 0.43 (0.02) and 3.65 (0.20) per 100.000 inhabitants in 1984 and in 1986, respectively. In 1991, the number of men (women) reported cases of AIDS reached 30.50 (5.95) cases per 100.000 inhabitants.

In addition, Brazil was strongly affected by HIV/AIDS from the mid-1980s (see Panel (a), Figure 1 and 2). The first outbreaks of the disease were detected on the coast of the southeast region, the most populous region of the country.[4] At the end of the 1990s, almost 148,000 people of both genders aged between 20 to 39 years were infected and the disease appeared in almost all Brazilian regions.[5] The disease spread so rapidly in the country that the Brazilian government was the first to provide free antiretroviral drugs to fight HIV/AIDS, starting in 1996, which decreased considerably the mortality of the disease (Dourado et al., 2006). In this sense, we explore how individuals behaved before and after the treatment became available.

To do that, we construct an annual panel of municipalities using data from the Brazilian Census, marital registries, and HIV/AIDS cases in two different time spans, 1982-1991 (Phase 1) and 1999-2010 (Phase 2).[6] We consider the single and marriage markets for individuals aged 20 to 39 years (young adults). We investigate if marital surplus increased at the start of the epidemic (1982-1991) and if this change in the surplus was maintained after the introduction of public policies that reduced the mortality risk of the disease. Our estimation has also two steps. First, we estimate a static marriage matching function following Choo and Siow (2006) and McFadden (1973), assuming that men and women have different traits (levels of education). The marital surplus is, therefore, our dependent variable. Second, we estimate a dynamic marriage matching function following Choo (2015), assuming that men and women have different levels of education.[7]

In our empirical model, we need to address potential endogeneity issues related to reverse causality, endogeneous income responses and other determinants of marriage changing over time. Then, besides including a set of relevant fixed effects and trends in our panel data model, we also construct an instrumental variable based on the HIV/AIDS infection of individuals

[4]The first wave of HIV/AIDS cases were mostly linked to male homosexuals and bisexuals, but it spread rapidly to heterosexual men and women aged 15 to 49. As shown in Panel (b) of Figure (1), heterosexual HIV/AIDS transmission grew rapidly from 1987 between men and women and faster for women from 1990. In 1984, the male/female ratio was 27/1, while in 1987 it was 9/1 and in 1991 it dropped to 4.9/1 (between men and women aged 20-39 years).

[5]In the state of Sao Paulo, the biggest state of Brazil and one of the most affected by the disease, reported cases in women (10 years or older) according to marital status, between January 1983 and June 1992, 39% were single, 27% were married, 13.7% were divorced or widowed and 20% had no information. In addition, 54% of married women were infected by sexual relations, 13% by drug use and 15% by drug use and sex, while in single women these percentages were 34%, 28% and 27%, respectively (Parker and Galvão 1996).

[6]The Census of 1991 included a question on the year of first marriage. To analyze the marital surplus in the second time span, we used marital registry data, only available by municipality from 1999 on. The marital surplus for the second period is calculated using the Brazilian Censuses of 2000 and 2010 and the civil registry data, both collected by the Brazilian Institute of Geography and Statistics (IBGE).

[7]See Chiappori et al. (2017) for an application of this method to calculate the marriage surplus. Bruze et al. (2015) also develop and estimate dynamic marital surplus.

from outside the marriage market of young adults (i.e., individuals between 65 and 85 years). We use the lags in t-1 and t-2 of this variable to avoid contemporary confounders. We argue that elderly individuals are out of the marriage markets of young adults and therefore there is no influence from the potential changes of former HIV/AIDS cases among elderly on labor markets for young adults.

Our findings suggest that marriage was used as a response to HIV/AIDS only in Phase 1 of the epidemic, i.e., before the start of the free treatment from the government. Our results wane when the mortality risk of the disease decreases in the second period. These results suggest that men and women changed their risk perception and sexual behavior when the disease started having effective and available treatment.[8] The estimates are robust to changes in the age ranges of the instrument and do not seem to be driven by changes in health infrastructure or income over time. The decision for marriage increased the marital surplus by 0.35% per additional HIV/AIDS case (out of 100,000 people) from 1982 to 1991 among people aged 20 to 39 years (Table 2).[9] The results from the empirical analysis are in line with what we find in the theoretical framework. After the emergence of the HIV/AIDS epidemic in the 80s, there was an increase in the marital surplus, which means the marriage was more valuable than the outside option (remaining single). We estimate a dynamic marital surplus, as in the theoretical framework, and find an increase of 0.30% per additional HIV/AIDS case (out of 100,000 people) from 1982 to 1991 among people aged 20 to 39 years.

Couples with high education drive our findings. The marriage surplus increased by 0.58% among most educated men and women. High-educated people were the first group (considering levels of education) affected by the epidemic (Brito et al., 2001). In addition, part of the literature shows that high-educated individuals have safer attitudes toward HIV/AIDS prevention (Duflo et al., 2006; Gao et al., 2012). We also find different responses for men and women when facing HIV/AIDS cases at the municipal level, with women receiving more from the increase in the share of the marital surplus. Our results suggest that women were more valuable in the marriage market, specially among the most educated. There is some evidence that women have less risky behavior than men, however, as the disease affected more men and the most educated, this may have affected their sexual behavior in response to the HIV/AIDS epidemic, increasing their search for more stable relationships and women's price

[8] Hart and Williamson (2005) showed that the risky sexual behavior of homosexuals in Scotland increased from 1999 to 2002, and it was correlated with the development of new therapies. Greenwood et al. (2019) argued that introducing policies that reduce the transmission risk can lead to a change in behavior, smoothing the policy's effect.

[9] We consider the database of single and married people, not including cohabiting.

in the marriage market (Moatti et al., 1991; Cohen and Bruce, 1997; Francis, 2008; Fernandez et al., 2019; Tenkorang et al., 2011; Greenwood et al., 2017; Spencer, 2020).

Our paper contributes to the scant literature on the relationship between the epidemic of HIV/AIDS (and other sexually transmitted diseases) and marriage empirically and theoretically. This literature is mostly based on randomized control trials in Africa (Dupas, 2011; Greenwood et al., 2017)[10], theoretical models to investigate marriage as a safeguard to HIV/AIDS (Angelucci and Bennett, 2021),[11] and models of sexual behavior such as Greenwood et al. (2019). Spencer (2020) shows that the birth rate among adult women in the US increased due to changes in sexual behavior (because of HIV/AIDS) and the use of contraceptive methods that reduced the likelihood of HIV/AIDS transmission.[12] Her analysis suggests that to avoid HIV/AIDS, women are shifting to monogamous relationships.

We also contribute to the empirical literature of marriage markets in three ways. First, we propose a quasi-experimental analysis which is representative of Brazil, a developing country with continental size. We estimate a static and a dynamic marital surplus. Second, we also investigate the marital surplus before and after the free distribution of antiretroviral drugs by the government. In this sense, we analyze how agents' incentives regarding the decision on marriage changed as the disease status changed from a death sentence to a chronic disease. Hence, our analysis provides a way of measuring how people respond to non-monetary incentives in relation to their marriage decisions before and after public policies that generate positive externalities. Finally, we also add to the theoretical framework by exploring the channel of sex behavior in response to increase in the probability in dating someone infected by HIV/AIDS.[13]

We add to the literature that shows that the reduction of expected cost from infectious diseases disincentivize protection measures (Posner et al., 1993; Laxminarayan and Malani, 2006; Hauck, 2018). This risk-elastic behavior was also found by Moatti et al. (1991) (when analyzing condom usage), and Geoffard and Philipson (1997); Tenkorang et al. (2011) (for other safe sex measures). In particular, Geoffard and Philipson (1997) and Hart and Williamson

[10] Dupas (2011) found that teenage girls react to information about HIV/AIDS, and Greenwood et al. (2017) showed that marriage reduced HIV/AIDS prevalence rates.

[11] Angelucci and Bennett (2021) constructed a model of assortative matching with two traits, one observable, attractiveness, and the other, sexual safety, hidden. They predicted that removing the asymmetric information about sexual safety changed the timing of decisions about marriage and pregnancy for safe respondents, and even more if they had the other attribute (attractiveness). To remove asymmetric information, in this case, meant increasing the frequency of HIV testing, which can allow safe people to signal and screen, increasing the probabilities of marriage and fertility.

[12] In addition, the epidemic also affected negatively the gonorrhea rate.

[13] The theoretical model focus on the first phase studied in this paper (1982-1991).

(2005) also investigated the change in risk behavior and ARV treatment beliefs and found evidence of the Pelztman effect (Peltzman, 1975), whereby individuals are less engaged in protective behaviors once the treatment is known. Kremer (1996) showed that the decrease in the probability of transmission can increase the prevalence rate because agents take more sexual risks. Interestingly, some studies found that women and men have different behaviors and risk perceptions (Moatti et al., 1991; Francis, 2008; Tenkorang et al., 2011; Cohen and Bruce, 1997).[14] Finally, this paper also adds to the literature that investigates the effect of how other shocks (income shocks) influence marriage (Hankins and Hoekstra, 2011; Corno et al., 2020).

The paper is organized as follows, besides this introduction. In Section 2 we briefly describe the fight against HIV/AIDS in Brazil. Section 3 presents the theoretical model that motivates our analysis. In the fourth section we present the identification strategy for the empirical analysis. Next, we present the marriage matching function that was used to calculate our measure of marital surplus. In Section 7 we describe the main results followed by robustness analysis. Finally, in Section 8 we conclude.

2 HIV/AIDS in Brazil

In 1982, Brazil had its first reported case of HIV/AIDS. The infection rates climbed fast afterwards, and in 1991 more than 8,000 new cases of people aged between 20 and 39 years were reported in the country (see Figure 1, Panel (a)). The Southeast region of the country was the most affected. The disease rapidly spread from homosexual and bisexual men to heterosexual men and women, from high educated to low educated individuals and among all income classes (see Figure 1, Panel (b)), as well as expanded from the Southeast (most developed region) to the rest of the country (Figure 2).[15] Panels (a) and (b) of Figure (1) also show that the number of HIV/AIDS infected individuals was higher for men. However, after 1990 this trend changed and more women were infected compared to men. And this trend continued until a few years ago. We observe in panel (c) of Figure (1) that after 1996, the number of HIV/AIDS deaths caused by the disease decreased in all Brazilian regions, especially in the Southeast region, where the disease had a higher prevalence in almost all periods.

[14] See Guillon and Thuilliez (2015) for a complete literature review.

[15] See Bacon et al. (2004), Szwarcwald et al. (1998), Brito et al. (2001) and Barbosa et al. (2009) for descriptions of the spread of the disease in Brazil.

[insert Figure (1) here]

The Brazilian government's response to the epidemic improved in the 1990s. From 1992, using both tax revenue and World Bank loans, the government started to finance the cost of HIV/AIDS treatment. The country forced the pharmaceutical companies to produce generic versions[16] of the ARV drugs. The ARV therapy involves a combination of HIV medicines that should be taken daily. They do not cure the infected individuals, but instead help them live longer, and reduce the risk of HIV transmission. The medicines reduce the amount of HIV in the body (to an undetectable level), consequently strengthening the immune system to fight off other infections[17]. Mainly due to this policy, Brazil is still globally recognized for its universal and free of charge provision of ARVs for people infected by HIV since 1996. In 1999, 85 thousand people received ARV treatment in Brazil and the number increased to 404 thousand people receiving the treatment in 2014. According to Brito et al. (2015), the government spent 400 million dollars for drug provision that year.[18]

Together with the drug distribution, the free provision of condoms and the awareness campaigns of condoms' importance were also important policies that have helped to control HIV/AIDS in the country since 1994. Information on disease transmission was also conveyed to sex workers (Dourado et al., 2015). Another relevant intervention was the needle exchange program, which stimulated drug users not to share needles (Levi and Vitória, 2002). Since 2005, the Brazilian Ministry of Health has also been providing rapid tests to diagnose HIV/AIDS.

Due to the combined policies to fight against HIV/AIDS, which had strong civil society participation, HIV/AIDS-related deaths decreased sharply after 1996 (Figure 1), especially in the Southeast region, where most cases were recorded. However, from 2000 onward there is an upward trend in the growth of deaths in some regions.

In 2017, there were 860 thousand Brazilians living with HIV/AIDS (13,000 of them were children aged 0 to 14), and 60% of the infected adults were receiving antiretroviral treatment (Avert, 2018). The fight against HIV/AIDS is still active in Brazil, mainly focused on higher risk groups (such as young homosexual men), where the rate of infection tripled from 2006-

[16] Generic drugs are unbranded pharmaceuticals, marketed under their active ingredient name without advertising.

[17] Information available at `https://hivinfo.nih.gov/understanding-hiv/fact-sheets/hiv-treatment-basics`. Individuals should start taking the medicines as soon as possible after the diagnosis.

[18] More information on the Brazilian policy against AIDS can be found in Levi and Vitória (2002) and Dourado et al. (2006).

2015.[19] Since 2017, Brazilian public health service has also been offering Truvada, which reduces the risk of contracting HIV/AIDS when taken daily, to higher risk groups (men who have sex with men, transgenders, and prostitutes) (Darlington, 2017).

3 Theoretical Motivation

To evaluate the decision between dating and getting married or remaining single and the increased likelihood of meeting someone infected with AIDS, we perform a two-period non-cooperative model with incomplete information. We intend to rationalize the decision about marriage from the early 80's to the beginning of the 90's (the first phase of our empirical estimation) and to capture the mechanisms by which the onset of the disease affected the payoff of marriage.

As the disease began to spread in the country from 1984 onward, the number of marriages and cohabitation rose 18% (per 100,000 inhabitants) in two years, as shown in Figure 1 (C). From 1986, there is a stabilization in the number of marriages and cohabitation, but a drop in the number of marriages until 1991. There is a substitution between formal marriage and cohabitation from 1987 to 1991. The main goal of this section is to rationalize the relationship between the rise in marriages and cohabitation, and the spread of the AIDS epidemic in Brazil (Figure 1 A).[20]

In the model, the decision to marry in the first or second period may change as the probability of dating someone infected with AIDS increases, considering that the individual's disease status is private information.[21] The findings in this section explain the mechanism we find in the empirical results, in which there is an increase in the marital surplus in the first period in relation to postponing the decision to marry. These results are shown in section (7.2) where we run a dynamic exercise.

3.1 Model with incomplete information

Consider the case where there are two equally large populations of men and women to be matched and the only difference between them is a type, high (H) or low (L), which means

[19] Recently, Brazil – together with other Latin American countries – established the "90/90/90" targets: 90% of people living with HIV/AIDS knowing their HIV status; 90% of them receiving antiretroviral therapy; and 90% of people on treatment virally suppressed by the year 2020 (Brito et al. (2015)).

[20] This model is limited to rationalize the effects of the disease in the 80s. The first phase presented in the empirical analysis.

[21] The baseline model with complete information is presented in the Appendix B.

they are healthy or have been infected by HIV. This is a private information. People know their status but the only information people have about others is the likelihood, $\phi_t(X)$, of meeting someone with HIV/AIDS in each period and this probability increases over time, that is, $\phi_2(X) > \phi_1(X)$.

Man (woman) decides to enter or not in a dating relationship, then they decide to get married or not in one of the two periods. Since we consider the payoff of entering a dating relationship is greater than remaining single, the agent decides to enter such a relationship. After that, nature plays and decides with probability p if the date is the right woman (man) and with probability (*1-p*) if the date is not the right person. The player observes the result and decides if he (she) will get married or not. If the person gets married in the first period, he (she) will receive a payoff in both periods, but if he (she) chooses to marry in the second, he (she) earns a payoff only in this period.

In the latter case, he (she) decides to engage in another dating relationship and he (she) observes the result after nature plays again, where with probability p the new date is the right woman (man) or with probability (*1-p*) she (he) is not. In this last period he (she) remains single and earns zero or gets married and earns a positive payoff. As the expected payoff is greater than zero, he (she) will always decide to get married instead remaining single. In this exercise, we do not include divorce or death rates similar to Angelucci and Bennett (2021).

The discount factor of a person i, δ_i, is idiosyncratic and distributed uniformly between 0 and 1: U $\sim$ [0,1]. $\phi_t(X)$ is the probability of dating a person infected by HIV (low type) during his (her) life. The random variable X, which denotes this probability, has a Bernoulli distribution played twice.[22]

The payoffs differ between types: Π^{hh} (both High types), $\Pi^{hl} = \Pi^{lh}$ (High type and Low type) and Π^{ll} (both Low type). We consider $\Pi^{hh} > \Pi^{hl} = \Pi^{lh} > \Pi^{ll} > 0$, which means that people prefer to get married rather than stay single. The values of p, x, Π^{hh}, Π^{lh}, Π^{ll}, $\phi_1(X)$ and $\phi_2(X)$ are common knowledge but it is private information if the person i is infected or not.

We consider that the probability of meeting someone infected by the disease is different between the two periods, and increases over time. As the proportion of infected people increases, men and women advance their decision regarding marriage. This result is shown in section 3.1.1. There is some evidence that men and women response to the HIV/AIDS epi-

[22] $\phi_t(X) = q^x(1-q)^{1-x}$, for x=0,1 and t=1,2. In this distribution, x is an integer, q is the chance of meeting person from the Low type (infected by HIV) and $1 - q$ is the chance of meeting someone of the High type. The probability q takes a value between $0 \leq q \leq 1$, and x has only two outcomes, success (1) or failure (0).

demic by changing their sexual behavior. Angelucci and Bennett (2021) shows that removing the asymmetric information (adverse selection) about sexual safety changes the timing about marriage and pregnancy for safe respondents. Spencer (2020), who finds an increase in birth rate due to women's increased monogamous relationships because of the emergence of the HIV/AIDS epidemic.

There is some evidence that women have less risky behavior than men, however, as the disease affected more men and the most educated, this may have affected their sexual behavior in response to the HIV/AIDS epidemic, increasing their search for more stable relationships and women's price in the marriage market

Equations (1) to (4) present the payoffs received in both periods. Equation (1) shows the expected payoff of people of high (low) type when nature chooses that the date is the right person:

$$E(\Pi_i^j|t=1) = (1+\delta_i).p.[\Pi_r^{jh}.(1-\phi_1(X)) + \Pi_r^{jl}.\phi_1(X)], \quad j=\{H,L\} \quad r=right \tag{1}$$

while the expected payoff of people of high (low) type when nature chooses that the date is the right one and the person get married in the second:

$$E(\Pi_i^j|t=2) = \delta_i.p.\{[p.[(1-\phi_2(X)).\Pi_r^{jh} + \phi_2(X).\Pi_r^{jl})] + (1-p).[(1-\phi_2(X)).\Pi_w^{jh} + \phi_2(X).\Pi_w^{jl})]\} \quad j=\{H,L\} \quad w=wrong \tag{2}$$

The expected payoff of high (low) type people when nature chooses the date is the wrong person and a man (woman) decides to engage in a new dating relationship is:

$$E(\Pi_i^j|t=2) = (1-p).\delta_i.\{p.[(1-\phi_2(X)).\Pi_r^{jh} + \phi_2(X).\Pi_r^{jl})] + (1-p).[(1-\phi_2(X)).\Pi_w^{jh} + \phi_2(X).\Pi_w^{jl})\} \quad j=\{H,L\} \tag{3}$$

Finally, the expected payoff of high (low) type people when nature chooses the date is the wrong person and man (woman) decides to marry in the first period:

$$E(\Pi_i^j|t=1) = (1-p).(1+\delta_i).[\Pi_r^{jh}.(1-\phi_1(X)) + \Pi_r^{jl}.\phi_1(X)] \quad j = \{H, L\} \tag{4}$$

Depending on the expected payoffs and the idiosyncratic discount factor, the agents decide when to get married.[23]

3.1.1 Why get married earlier?

The hypothesis we consider in this framework is that the proportion of people who decide to get married in the first period increases, advancing their decisions on marriage, as the proportion of people infected with HIV/AIDS grows (the likelihood $\phi_2(X) > \phi_1(X)$). Consider $\phi_2(X) = \phi_1(X) + \epsilon$, where $\epsilon > 0$. We want to show that an increase in ϵ increases the marital surplus in the first period, which also increases the proportion of people who get married earlier. Using equations (1) and (2) and replacing $\phi_2(X)$ by $\phi_1(X) + \epsilon$, we analyze two cases:

Case 1: We calculate the difference between the two terms (equation (5)), $\Upsilon(\pi_i^h)$, which is the expected payoff of people not infected by the virus who get married when nature chooses the date is the right person in the first period minus the expected payoff when the person enters a new dating relationship and gets married in the second. We want to study the impacts on marital surplus in the first period caused by changes in ϵ.

$$\begin{aligned} \Upsilon(\pi_i^h) = (1+\delta_i).p.[\Pi_r^{hh}.(1-\phi_1(X)) + \Pi_r^{hl}.\phi_1(X)] - \\ p.\delta_i.\{[p.[(1-(\phi_1(X)+\epsilon)).\Pi_r^{hh} + (\phi_1(X)+\epsilon).\Pi_r^{hl})] + \\ (1-p).[(1-(\phi_1(X)+\epsilon)).\Pi_w^{hh} + (\phi_1(X)+\epsilon).\Pi_w^{hl})]\} = 0 \end{aligned} \tag{5}$$

Proposition 1 (The role of the likelihood $\phi_i(X)$ in the expected payoff when nature chooses the date is the right person)**.** *Under the assumptions $\phi_2(X) > \phi_1(X)$, $\Pi^{hh} > \Pi^{hl} = \Pi^{lh} > \Pi^{ll} > 0$, the values of p, x, Π^{hh}, Π^{lh}, Π^{ll}, $\phi_1(X)$ and $\phi_2(X)$ are common knowledge but compose private information if the person i is infected, it is possible to show that the increase in the probability of dating someone infected by AIDS increases the expected payoff in the first period compared to the second. It is a positive function of the parameter ϵ, $\frac{\partial \Upsilon}{\partial \epsilon} > 0$. The same holds for Low-type people.*

[23] Full details of the decision on marriage is shown in Appendix C.

Proof. Calculating the derivative of equation (5) in relation to the ϵ parameter, we have a positive sign. This indicates that an increase in ϵ increases the value of Υ, that is, there is an increase in the marital payoff of the first period in relation to the second.

$$\frac{\partial \Upsilon}{\partial \epsilon} = -p.\delta_i.\{p[\Pi_r^{hl} - \Pi_r^{hh}] + (1-p)[(\Pi_w^{hl} - \Pi_w^{hh})]\} > 0 \tag{6}$$

As a consequence, a marginal increase in ϵ affects the proportion of men (women) who decide to marry in the first period relative to the group who get married in the second, since ϵ increases the payoff in the first period relative to the second. The same result holds for low-type people. □

Case 2: We calculate the difference between the two terms as in equation 7 (difference between equations ((3) and (4)), function $\Lambda(\pi_i^h)$, which is the expected payoff of a person not infected by the virus when nature chooses that the date is the wrong person in the first period and of getting married in the second minus the expected payoff when the person gets married in the first period. Again, we want to study the impacts on marriage caused by changes in ϵ.

$$\begin{aligned} \Lambda(\pi_i^h) = (1-p).\delta_i.\{p.[(1-(\phi_1(X)+\epsilon)).\Pi_r^{hh} + \\ (\phi_1(X)+\epsilon).\Pi_r^{hl})] + (1-p).[(1-(\phi_1(X)+\epsilon)).\Pi_w^{hh} + (\phi_1(X)+\epsilon).\Pi_w^{hl})\} - \\ (1-p).(1+\delta_i).[(1-\phi_1(X)).\Pi_w^{hh} + \phi_1(X).\Pi_w^{hl})] = 0 \end{aligned} \tag{7}$$

Proposition 2 (The role of the likelihood $\phi_i(X)$ in the expected payoff when nature chooses the date is the wrong person)**.** *Under the same assumptions as in Proposition 1, it is possible to show that the increase in the probability of dating someone infected by AIDS decreases the expected payoff in the second period compared to the first. It is a negative function of the parameter* ϵ*,* $\frac{\partial \Lambda}{\partial \epsilon} < 0$*. The same holds for Low-type people.*

Proof. Calculating the derivative of equation (7) in relation to the ϵ parameter, we have a negative sign. This indicates that an increase in ϵ decreases the value of Λ, that is, there is an increase in the marital payoff of the first period relative to the second.

$$\frac{\partial \Lambda}{\partial \epsilon} = -(1-p).\delta_i\{[\Pi_r^{hh} - \Pi_r^{hl}] + [\Pi_w^{hh} - \Pi_w^{hl}]\} < 0 \quad (8)$$

As a consequence, a marginal increase in ϵ affects the proportion of people who decide to marry in the second period relative to the group who get married in the first. As ϵ increases, the payoff in the second period decreases relative to the first. The same result holds for infected people. □

Therefore, the increase in the likelihood of meeting someone infected by AIDS affects the expected payoff of getting married in first period relative to the second one, making it more attractive to marry in the first rather than the second. This is the movement we observe in Panel (c) of Figure (1), where the number of marriages and cohabitation increases after the onset of the AIDS epidemic in 1984. Hence, the increased risk of becoming infected with the disease may explain why we found in the empirical analysis (section 7.2) an increase in marital surplus compared to being single in a dynamic estimation.

4 Empirical Strategy

We established a potential relationship between the HIV/AIDS epidemic and marriage, but we need to investigate what are the possible reasons for this relationship. We argue that the main channel is risk perception and its effects on sexual behavior.

The risk perception link between HIV/AIDS epidemic and marriages rely on the literature that relates risk perception with sexual behavior (Spencer, 2020; Greenwood et al., 2017). Before the introduction of ARVs (antiretroviral drugs), the disease was a death sentence. The disease was spread from homosexual and bisexual men to heterosexual men and women mainly through sexual intercourse. The inherent risk of catching the disease was more significant for men at the beginning of the epidemic. The perception of risk can lead more men (compared to women) to seek fewer partners, moving towards monogamous and more stable relationships, which would increase the number of marriages. Moreover, we believe this is the reason why we find heterogeneity in the results where women's marital surplus increased more than men's (Panels (a), (b) and (d) of Figure (3)). Evidence that women are more valued in the marriage market. Our theoretical framework also discusses this mechanism showing that an increase in the likelihood of being infected by the disease raises the expected payoff of marital surplus in the first time span (1982-1991).

After the introduction of the ARV treatment, individuals may engage in less protective behavior as treatment is more efficient.[24] This behavior is in line with the results of Hart and Williamson (2005), and Kremer (1996) and also with our findings that after the introduction of free distribution of ARVs the effect disappeared.

To disentangle the effect of HIV/AIDS on adult marriages, we need to address at least three potential endogeneity issues. First, an endogeneous response can occur because the epidemic might also have affected adults' labor productivity, and hence their income (Laxminarayan and Malani, 2006; Hauck, 2018). This indirect effect of HIV/AIDS on income might bias our results, since income can be an important determinant of marriage. Thus, disregarding the income channel of the epidemic would induce an upward bias in the estimated impacts of HIV/AIDS on marriage. Second, there might be other determinants of marriage that changed over time and are correlated with the evolution of cases of HIV/AIDS. One example is the change in individuals' behavior during the 1980s and 90s. The increase of sexual liberalization of the period and the change in women's behavior regarding marriage might be timely correlated with infection by HIV/AIDS. This other potential channel of the epidemics would induce a downward bias in the estimated impacts of HIV/AIDS on marriage. In addition, there is a possibility of reverse causality, where marriage can affect the transmission of HIV/AIDS. All these potential biases would reduce the likelihood of consistently estimating the effect that HIV/AIDS can have on marriage.

To deal with the potential biases, we propose an instrumental variable approach. The exclusion restriction states that we need one (or more) instrument that (i) is not a direct determinant of marriage of individuals between 20 and 39 years and that do not correlate with income generated by those adults, or with the younger men and women's decision over time, and (ii) is strongly correlated with the regional spread of HIV/AIDS.

We use the lags of HIV/AIDS rate among elderly people (above 65 years) as instruments. Former HIV/AIDS cases among elderly people is expected to be correlated with HIV/AIDS cases among individuals between 20 and 39 years due to the spread of the disease across the Brazilian municipalities. However, we argue that HIV/AIDS cases among older individuals in the past do not affect the marriage market of younger adults (and the size of the marriage market)[25] nor their income.

Since older individuals are mostly not in the labor force, we posit that adult incomes are not

[24] See Panel (c) of Figure (3)

[25] We use lags because the increase in marriages might negatively affect the HIV/AIDS transmission between all age groups. Therefore, the cases in previous years prevent us to include this potential endogeneity.

affected. As we also construct our marital surplus considering married and single individuals only between 20 and 39 years, we also hypothesise that the HIV/AIDS rate among elderly people have not determined the marriage markets of young adults. Similar instruments were also proposed by the literature, such as Grimard and Parent (2007) and Rodríguez-Planas and Tanaka (2022) that use as instruments information for different birth cohorts to understand the impact of education on smoking behavior Grimard and Parent (2007), and of gender norms on women's labor participation Rodríguez-Planas and Tanaka (2022).

Therefore, we estimate our model in two stages. First, we regress HIV/AIDS cases among young adults ($AIDSy$) on the disease cases among the elderly population ($AIDSe$), with fixed effects and controls. Then, we regress marital surplus, from the static and dynamic models, on the fitted cases of HIV/AIDS from the first stage, correcting the standard errors by clusters. All regressions include the same set of controls and fixed effects, as well as population weights, according to the following equations:

$$AIDSy_{mt} = \alpha_t + \alpha_m + \alpha_s t + \rho_1 AIDSe_{mt-1} + \rho_2 AIDSe_{mt-2} + \beta X_{mt} + \varsigma_{mt} \tag{9}$$

$$Marriage_{mt} = \theta_m + \theta_t + \theta_s t + \ \tau \widehat{AIDS} y_{mt} + \beta X_{mt} + \varepsilon_{mt} \tag{10}$$

where m denotes the Brazilian municipalities, and t years from 1982 to 1991 (Phase 1) and 1999 to 2010 (Phase 2); $Marriage$ represents the marital surplus, which will be discussed in Section 5; θ_m is the municipality fixed effect; θ_t is year fixed effect; $\theta_s t$ represents a state-specific trends; $\mathbf{X}$ is a vector of controls, including the mean age and mean years of schooling of the sample by municipality; $AIDSy_{mt}$ denotes the reported cases of HIV/AIDS per 100,000 population by municipality in year t for individuals aged between 20 and 39 years, $AIDSe_{mt-1}$ the reported cases of HIV/AIDS per 100,000 population by municipality in year $t-1$ for individuals older than 65 years and $AIDSe_{mt-2}$ the reported cases of HIV/AIDS per 100,000 population by municipality in year $t-2$ for individuals older than 65 years.[26]

The variables of the dynamic specifications are constructed in the following way. The regression is estimated every two years (1982,1984,1986,1988 and 1990). We consider as independent variable the number of cases of AIDS for years 1982,1984,1986,1988 and 1990. We consider the number of cases of AIDS for individuals older than 65 years in $t-2$ as instrument. We add the municipality and year fixed effects.[27]

[26]We present some robustness results of the instrument by widening and tightening the age range of the HIV/AIDS's incidence (more than 55 years, 60 years, or 65 years). The results are qualitatively the same.

[27]Details of how the variables are constructed are in Appendix A.

5 The Marital Surplus

In this section, we describe how we estimate the marital surplus, considering that men and women have different traits. In the first part, we follow (Choo and Siow, 2006) and (McFadden, 1973) to estimate the static marital surplus. In the second, we follow (Choo, 2015) to estimate the dynamic marital surplus.

5.1 The Static Marital Surplus

Consider a large number of I types of men and J types of women where types can be specified according to level of schooling, age, religion, race, ethnicity etc. Men and women can choose between getting married or remaining single. The person each man or woman will marry depend on the preference parameters I (J). In the data it is possible to observe each type m_i of men in the marriage market and the same for women (w_i).

A marriage market function is defined as an I x J matrix $\mu(M, W; \Pi)$. This function must satisfy the following constraints as shown in Choo and Siow (2006):[28]

$$\mu_{0j} + \sum_{i=1}^{I} \mu_{ij} = w_j \quad \forall j, \tag{11}$$

$$\mu_{i0} + \sum_{j=1}^{J} \mu_{ij} = m_i \quad \forall i, \tag{12}$$

$$\mu_{i0} + \mu_{0j} + \mu_{ij} \geqslant 0 \quad \forall i, j. \tag{13}$$

where μ_{ij} is the number of men of type i married to women of type j, μ_{i0} is the number of unmarried men of type i and μ_{0j} is the number of single women of type j.

In the framework of transferable utility, a man of type i who marries a j type of woman must transfer to her an amount of income represented by τ_{ij}. In this market, we shall have I x J marriages. This market clears when, given the equilibrium transfers τ_{ij}, the supply and demand of men and women are equal, that is, the number of men of type i equals the number of women of type j and vice-versa, considering all i, j.

In order to estimate the model described above we will follow Choo and Siow (2006). Men and women maximize their marriage surplus by observing that their peers are doing likewise.

[28]The vector M composed of available men with types m_i for all $i = 1, ..., I$ and W represents the vector composed of available women with types w_j for $j = 1, ..., J$.

The problem of linear programming developed by Shapley and Shubik (1971) posits that any stable matching between men and women emerges from the maximization of the aggregate of marriage surplus considering all possible assignments. The shares that women and men receive are determined endogenously.

Consider μ_{ij}^d the number of marriages demanded by type i men with type j women, μ_{i0}^d the number of single type i men, α_{ij} the share a man receives from the systematic return from marriage and τ_{ij} is the transfer made by man to woman.

$$\ln \mu_{ij}^d = \ln \mu_{i0}^d + \alpha_{ij} - \tau_{ij} \tag{14}$$

Equation (14) is the quasi demand equation for men of type i who are married to j type women. In turn, equation (15) is the quasi-supply equation of women of type j who are married to men of type i and γ_{ij} the share a woman receives from the systematic return from marriage.

$$\ln \mu_{ij}^s = \ln \mu_{0j}^s + \gamma_{ij} + \tau_{ij} \tag{15}$$

Considering that the market clears when $\mu_{ij}^s = \mu_{ij}^d$ and from equations (14) and (15), we have the following result:

$$\ln \mu_{ij} - \frac{\ln \mu_{i0} + \ln \mu_{0j}}{2} = \frac{\alpha_{ij} + \gamma_{ij}}{2} \tag{16}$$

If we consider $\pi_{ij} = \ln \Pi_{ij} = \frac{\alpha_{ij}+\gamma_{ij}}{2}$, the equation (16) can be written as:

$$\Pi_{ij} = \frac{\mu_{ij}}{\sqrt{\mu_{i0}\mu_{0j}}} \tag{17}$$

The equation above is the marriage matching function. Taking the log of the right side of the equation (17) gives the marriage surplus of the couple type i, j compared to the gain if each partner had remained single.[29]

Considering equations (14), (15) and π_{ij}, we can identify the shares received by men and women.

$$\ln \left(\frac{\mu_{ij}}{\mu_{i0}} \right) = \alpha_{ij} - \tau_{ij} = n_{ij} \tag{18}$$

[29] In the case there are more single men of type i and women of type j in the population compared to the number of marriage between i, j partners. By using geometric average in the number of unmarried partners of types i and j, it is possible to control for these effects.

$$\ln\left(\frac{\mu_{ij}}{\mu_{0j}}\right) = \gamma_{ij} + \tau_{ij} = N_{ij} \tag{19}$$

In this study, we calculate equations 17, 18 and 19 and then aggregate the results by municipality.

5.2 The Dynamic Marital Surplus

In the second part, we estimate the dynamic marital surplus following (Choo, 2015). We estimate the following equations:

$$\Pi_{ij} = \ln\left(\frac{p_{ij}.q_{ij}}{\prod_{t=0}^{T_{i,j}-1}(p_{i+t,0}.q_{0,j+t})^{(\beta^t)}}\right) \tag{20}$$

where $t = 1, ..., T$, $p_{ij} = \mu_{ij}/m_i$, $q_{ij} = \mu_{ij}/f_i$, β is the discount factor, m_i is the number of available men (married and single) and f_i is the number of available women (married and single).

$$n_{ij} = \ln\left(\frac{p_{ij}}{\prod_{t=0}^{T_{i,j}-1}(p_{i+t,0})^{(\beta^t)}}\right) \tag{21}$$

$$N_{ij} == \ln\left(\frac{q_{ij}}{\prod_{t=0}^{T_{i,j}-1}(q_{0,j+t})^{(\beta^t)}}\right) \tag{22}$$

Equation (20) estimates the couple's present discounted utility from being married today in a partnership (i, j) compared to the present discounted sum of the payoffs from being single for t periods. Equations (21) and (22) present the share of the marital surplus received by men of type i and women of type j married in a match (i, j), respectively. In this work we consider $t = 1, 2$. We estimate the marital surplus every two years. [30]

6 Data

To investigate if the relationships highlighted in Sections 3 and 4 are observed in the data, we use data from two different time spans (1982-1991 and 1999-2010). The idea is to test if the effect of HIV/AIDS spread in the country changed in relation to marriage decision with the

[30] More details about the variables are presented in section A.

introduction of the free distribution of antiretrovirals in 1996. We present the data separately since the available data changed over time.

6.0.1 Phase 1: HIV/AIDS rise (1982 to 1991)

To access data on marital surplus for the first period, we use the 1991 Brazilian Census, collected by the Brazilian Institute of Geography and Statistics (IBGE). The Census data include the year a person gets married for the first time. We use this variable to construct our dependent variable. Since we focus on marriage decision, we exclude widowed or divorced individuals from the sample. Thus, the sample is composed of men and women from 20-39 years who remained single (who never married and is not currently married), or married (for the first time only) between 1982 and 1991. We form two databases from this sample, the first one is the full sample itself, which is composed by first married, first cohabiting and single people.[31] The second database is composed only by first married people and single people. In order to estimate the effects by different types of individuals by levels of education we divide them in two groups: the first one composed by individuals with elementary and middle school and the second with individuals with high school or above. We use the second database in order to be able to compare this period (1982-1991) with the second one (1999-2010) where the available data are for married people only. In the Appendix A, we describe in more details how the variables are constructed.[32]

Between 1980 and 1991, more than 500 new municipalities were created in Brazil. Therefore, we use the minimum comparable areas (AMCs in the Portuguese initials) from 1970 to 2000, constructed by IBGE, which represent the most disaggregated regional unit to compare municipalities.[33] Thus, the municipalities are grouped by AMCs.

We use the number of reported cases of HIV/AIDS by municipality and year from the Information System for Notifiable Diseases (SINAN) of the Ministry of Health from 1980 to 1991.[34] Then, we calculate the number of cases per hundred thousand individuals using the total population broken down by age, municipality and year from IBGE. We calculate the HIV/AIDS cases per 100,000 for the following age ranges: 20 to 39 years, more than 55 years; more than 60 years; and more than 65 years. The other age groups are calculated because of

[31] In Brazil, the Constitution of 1988 established equal rights of cohabitation and marriage. Cohabitation in the 1980s was more prevalent among the poorer people.

[32] We use survey sampling weights to balance the data and calculate all variables by municipality.

[33] This variable is available at the website ibge.gov.br. There are 3,658 AMCs, while the number of municipalities in 1991 was 4,491.

[34] Data for 1981 is not available in the SINAN database (website consulted in December 2018:`http://www2.aids.gov.br/cgi/deftohtm.exe?tabnet/br.def`

the instruments and robustness analysis.

We also construct variables to control the regressions, or to understand the heterogeneity of the marriage response to the HIV/AIDS epidemic. We calculate the number of general hospital beds per population to investigate changes in health infrastructure over time. The number of hospital beds is available by municipality and year from the Medical-Sanitary Assistance Survey from 1981-1990 (except in 1988), conducted by the Ministry of Health.[35]

Second, to investigate if our instruments work, and therefore we are not capturing the income effect of the epidemic on marriage, we calculate the average income using RAIS (Annual List of Social Information) database. The variable is the mean income (in multiples of the minimum monthly wage) in logarithm using RAIS data aggregated by municipality and year from years 1985 to 1991 (they are available from 1985 on).[36] The age range used is 18 to 39 years.

6.0.2 Phase 2: Treatment for HIV/AIDS (1999 to 2010)

The marital surplus for the second period is calculated using the Brazilian Censuses for 2000 and 2010 and the civil registry data, both collected by the IBGE. We need to change the methodology to create the variable since the most recent censuses do not collect the year individuals got married for the first time (as was the case in the 1991 Census). In this sense, we use Census data to calculate the total single people by year and use data from civil registries by municipality and year to collect information about first marriages from 1999 to 2015. Again, we exclude widowed, widows, and divorced people from the sample. The final sample includes men and women from 20 to 39 years old who remained single (never married and is not currently married), or married (first unions only) between 1999 and 2010. We also analyze the marital surplus with two different types of men and women, considering different levels of education.[37]

We do not analyze the period from 1992 to 1998 since the marriage dataset is not available by municipality. We also do not include more recent years since data for single individuals are only available until 2010 (last population Census in Brazil). However, years from 1999 to 2010 are interesting since they are after the implementation of free distribution of ARVs in

[35] The data for 1988 and 1991 are not available due to change in the questionnaires (1988) and collection problems (1991). The website was consulted in September 2018 at http://tabnet.datasus.gov.br/cgi/deftohtm.exe?ams/cnv/amsobr.def

[36] We run robustness tests using the variable income.

[37] In the Appendix A, we describe in more detail how the variables are constructed. We follow the age division of the dataset.

Brazil (from 1996 on).[38]

The number of reported cases of HIV/AIDS by municipality and year are collected from the same source as the first period: Information System for Notifiable Diseases (SINAN) of the Ministry of Health from 1999 to 2010, and we construct the indicator according to population using IBGE data for all age ranges of the populations.

7 Results

7.1 Main results

We show our first stage results in Table 1 for both phases we analyze: Phase 1 (1982 to 1991), in Columns (1) and (2)[39]; and Phase 2 (1999 to 2010), in Column (3). We have included the incidence of HIV/AIDS in elderly in t-1 and t-2 as instruments. Our instruments are statistically significant in both Phases and the sign of the variables are as expected: the greater the incidence of HIV/AIDS cases in the elderly, the greater the incidence of the disease in the younger people. The instruments seem to be stronger for the first period of analysis (the F-statistics of the excluded instruments are higher than 36) than for the second period (F-statistics of 6.11).[40]

Then, we present the marital surplus regression [41] on cases of HIV/AIDS in young adults in Table 2. Column (1) presents the OLS regressions with fixed effects and controls (municipality and year fixed effects, state trends, average age of the married individuals). Column (2) displays the second stage results using as instruments the number of HIV/AIDS cases among the elderly in t-1, and t-2, besides all fixed effects and controls. Each panel refers to the results of a different period (Panels A and B show the estimated coefficients of Phase 1, while Panel C shows the estimates of Phase 2), and sample (Panel A considers the marital surplus from marriages and co-habitation, and Panels B and C consider only the marital surplus from marriages).

[insert Table (1) here]

[insert Table (2) here]

[38]http://www2.datasus.gov.br/DATASUS/index.php?area=0203&id=6930 and https://sidra.ibge.gov.br/acervo#/S/RC/A/4/T/Q.

[39]The sample in Columns (1) and (2) follows the same sample of municipality-year of the second stage.

[40]We did a robustness check in the second period adding the contemporary variable of elderly as instrument as shown in Appendix Table 5. The F-statistics raises from 6.11 to 24.6.

[41]The dependent variable is the marital surplus, calculated using Equation 17.

The turning point in the number of marriages and cohabitation in 1984 showing an increase in marriages as the incidence of the reported cases of AIDS rise, as indicated in Figure (1, Panel (c)) for the period 1982-1991, may explain our findings. After including controls, we find a positive relationship (Column (1), Panels A and B) between cases of AIDS and the marital surplus. When we estimate the same regression using instrumental variables, we find an even more positive effect of the increase of HIV/AIDS cases on the marriage surplus, indicating that we still have a negative bias, probably because of the changes in men's and women's behavior over time that are not perfectly measured by the time-varying controls and stat-trends that we included in the regression.

The HIV/AIDS effect on marital surplus is statistically significant at 1% and shows that total gain of marriage per partner increases by 0.22% by additional case of AIDS per 100,000 in the first period (Panel A, Column 2), indicating that the effects were relevant when the disease appeared in the 80s. The same effect is 0.35% when we consider only married people (Panel B, Column 2). The results are higher for married people compared to married and cohabiting people. This difference may be evidence that risk attitudes affect the timing of marriage, where risk averse people tend to marry early as presented in Spivey (2010) and Schmidt (2008).[42] In addition, these results are in line with the idea that individuals sexual behavior responded to incentives, as discussed by Cohen and Bruce (1997), Greenwood et al. (2017)), Spencer (2020) and Greenwood et al. (2019).

In Phase 2, after the introduction of ARVs, the perceived vulnerability of the disease decreased and long-term monogamous relationships are not affected by HIV/AIDS cases (Panel C, column 2 shows no statistically significant coefficient for the second period). Figure (1, Panel (d)) shows tha there is a downward trend in the number of marriages in 1999. This result is also in line with Geoffard and Philipson (1997) and Hart and Williamson (2005), who found that agents are less involved in protective measures once effective treatment is available.[43]

The same behavior is observed when we compare the marital gains for men and women separately (equations 18 and 19). In Phase 1, the marital surplus for women relative to being single increased 0.25% (0.39%), while for men it rose 0.18% (0.32%), as shown in Panel A (B), Columns (2) and (3), of Table 3.[44] All estimates were statistically significant at 1%

[42]Marriage and cohabitation rights were different until 1988. Since then, these differences in rights have diminished in Brazil.

[43] We check the coefficient in the second stage adding the contemporary variable of elderly people as instrument. The F-statistics raises (Appendix Table 5) but the coefficient remains not statistically significant.

[44]The first result in the marital surplus considers married and cohabiting people whereas the second, between

(Panels A and B, columns (1) and (2) and at 5% column (3). The results suggest that the effect of the HIV/AIDS epidemic in the marital surplus is stronger for women in the first period of analysis, affecting more positively their shares in the marital surplus. Again, we find no significant effect for the second phase we analyze (1999 to 2010) for both men and women separately (Panel C, Columns (2) and (3)). These results are in accordance with the development of the disease in Brazil. As men were firstly more affected by the epidemic and to avoid infection, they may have searched for more stable and monogamous relationships compared to women.[45] Hence, women's price in the marriage market increased more than of men's. Overall, the result was an increase in marital surplus, with a more pronounce raise in women's share than men's.

[insert Table (3) here]

We now investigate which individuals (and couples) drive the results we find in Phase 1. We split the sample of young adults into women and men types. To divide individuals by types, we consider two levels of education: Low (elementary and middle school) and High (high school or above). We then calculate the marital surplus for couples of all combinations of types: Low/Low (men and women of type low); Low/High (men of type low and women of type high); High/Low (men of type high and women of type low); and High/High (when both men and women are of type high). Individuals with more years of schooling were the first ones affected by the epidemic (Brito et al. (2001)). Table (4) presents the estimated coefficients for the total sample (Panel A), men (Panel B) and women (Panel C).

The couples with high education drive our findings, and women seem to have appropriated more of the marital surplus than men. This result is in line with the literature that Duflo et al. (2006) and Gao et al. (2012) show that education level explains many attitudes toward HIV/AIDS prevention and sexual behavior and also with the spread of the disease in Brazil according education levels.

7.2 Dynamic Main Results

The increase in the likelihood of meeting someone infected by AIDS affects the expected payoff of getting married in first period relative to the second one, making it more attractive to marry in the first rather than the second. This increased risk of becoming infected with

brackets, only married people.

[45] Cohen and Bruce (1997) observed that real risk and perceived vulnerability of HIV infection differs by gender.

the disease may explain why we found in the empirical analysis an increase in marital surplus compared to being single.

To understand how our results change when considering a dynamic model as in the theoretical section, we now run the same set of results, but calculating the marital surplus for a two-period model. The dynamic exercise compares the gain of getting married with the outside option of remaining single in periods t and t+1. In the theoretical model, as the transmission rate increases, this information is known, but at the individual level the information is private, and the agents' decision regarding marriage (and the decision to remain single) may change. This result is observed in the dynamic results. The decision on marriage increases as the disease spreads in the country. Therefore, individuals may prefer to marry earlier, especially those who belong to the most affected groups, such as men and the most educated. In this context, women benefit, as their share in marital surplus increases more than men's.

Table 5 shows the first stage results for both phases we analyze: Phase 1 (1982 to 1991), in Columns (1) and (2); and Phase 2 (1999 to 2010), in Column (3). We have included the incidence of HIV/AIDS in elderly in t-2 as instrument. For the dynamic case, we use the marital surplus data every two years. Thus, we only use one lag of the instrument. The instrument seem to work for the first phase of analysis (the F-statistics of the excluded instrument is higher than 6.8), but it does not work for the second period (F-statistics of 0.26).

[insert Table (6) here]

Table 6 reports the second stage results. All panels display the second stage results using as instrument the number of HIV/AIDS cases among the elderly in t-2, besides all fixed effects and controls. Each panel refers to the results of a different period (Panels A and B show the estimated coefficients of Phase 1, while Panel C shows the estimates of Phase 2), and sample (Panel A considers the marital surplus from marriages and cohabitation, and Panels B and C consider only the marital surplus from marriages). For the dynamic model, we also find a positive relationship (Column (1), Panels A and B) between the dynamic gains of marriage and cases of HIV/AIDS in the Phase 1. The HIV/AIDS effect on marital surplus (using the dynamic model) is now statistically significant at 5%, considering only married people (Panel B, Column 1), and shows that total gain of marriage per partner increases by 0.30% by additional case of AIDS per 100,000 in the first period (Panel A, Column 1), indicating that the effects were relevant when the disease appeared in the 80s.

Similar to the static model, the results from Columns (2) and (3) suggest that the effect of the HIV/AIDS epidemic in the marital surplus is stronger from women in the first period of analysis. It seems that their shares in the marital surplus increased more than men's. Again, we find no significant effect for the second phase we analyze (1999 to 2010) for both men and women separately (Panel C).

[insert Table (7) here]

Using the surplus calculated by the dynamic model, we also investigate what type of individuals (and couples) drive the results we find in Phase 1. We split the sample of young adults into women and men educational types. Table 7 presents the estimated coefficients for the total sample (Panel A), men (Panel B) and women (Panel C). The results indicate that the increase in marital surplus was given by the increase in women's share in it by 0.61%.

7.3 Robustness Checks

Change in instruments: We now check the robustness of our results to different age ranges of the instruments. We test our instruments by varying the age periods of elderly people. The first stage results are presented in Appendix Table 1. The instruments with different ages (+ 55, + 60 and + 65) are all statistically significant in the first period (1984-1991). In the second (1999-2010), only + 65 is statistically significant. This suggests a greater proximity to the independent variable, that is, cases of AIDS of people aged between 20 and 39 years. The second stage results of the marital surplus presented in Table 8 (first column (second)), for the period between 1982 and 1991, are positive 0.18%(0.23%), 0.22%(0.29%) and 0.22%(0.35%), for all different instruments, respectively. The results are also robust to the change in instruments for the both periods of analysis.

[insert Table (8) here]

Lagged periods: The idea of this robustness check is to test our model considering the independent variable with lags of one and two years, to see if there is a temporal correlation between cases of AIDS and marital surplus. The results are shown in Table 9 where we can see a temporal correlation between instrumentalized AIDS cases and marital surplus (only married people) in the first phase (1982-1991). The temporal correlation is stronger with one lag (statistically significant at 1%) and greater than the coefficient reported in Table 2 (0.61% to 0.35%). However, we do not find this temporal correlation in one lagged period for the second period estimation (1999-2010).

[insert Table (9) here]

Other marriage determinants - Hospital Beds: We now check different dependent variable to understand their potential relationship with cases of HIV/AIDS that could explain partially the effects we estimate. In Appendix Table 2, we use the same specification as Equation 10 but now using the number of hospital beds per 100,000 inhabitants as dependent variable instead of marital surplus. Panel A shows the estimates for the static model, while Panel B displays the same coefficients calculating surplus using the dynamic model. We find that cases of HIV/AIDS for young adults do not explain the availability of hospital beds at the municipal level, which indicates that capacity constraint was not affected by the increase of HIV/AIDS cases in the period.

Other marriage determinants - Income: We also perform the same test, but now using average income in the municipality (total and for men and women separately) as dependent variables to investigate if HIV/AIDS affected labor productivity in the first period. The income variables are in logarithm and results are shown in Appendix Table 3. We also find a non-significant effect of women's and men's income. When running our baseline results including income as control (we lose part of the sample because the income variable is only available from 1985 on), we find a negative but non-statistical effect of income on the marital surplus as shown in Appendix Table (4), and the effect of HIV/AIDS is robust to our baseline regression.

8 Final Remarks

The HIV/AIDS epidemic seems to have affected preferences for marriage among adults in Brazil, especially during the 1980s. Our results suggest that the epidemic increased the static and dynamic marital surplus. Probably, marriage was used as a response to the HIV/AIDS epidemic. At the beginning (in the 80's), men were the most affected group, and this may explain why women benefit more as their share in the marital surplus increased more than men's, and by more educated individuals.

We find that the impact of HIV/AIDS on marital gains was limited due to the easy effective treatment in Brazil. As a consequence, the reduction of expected cost from infectious diseases disincentivize protection measure, in this case, marriage. When individuals' perception about the disease risks changed, the impact of HIV/AIDS on marital gains changed as well. The increase of stable and monogamous relationships seemed to have worked as a defense against the disease. The result is especially interesting since the number of HIV/AIDS cases in Brazil is greater among men than women. This could be evidence why the response regarding women's marital surplus increased more than men's. In addition, the results shown in the theoretical motivation are in line with what we find in the empirical part.

Our results are in accordance with the literature that has found a relationship between marriage and HIV/AIDS, including that an increase in marriage can be seen as an imperfect protection against the virus. Our paper also contributes to the literature by measuring the role played by non-monetary incentives in marriage decisions, as well as how incentives changed with universal access to effective treatment.

Compliance with Ethical Standards

Funding: This study is based on Lorena Hakak's posdoctoral research at the University of Sao Paulo. She received financial support from the Office to Coordinate Improvement of University Personnel (CAPES) to pursue her postdoctoral studies.[46]. Paula Pereda also acknowledges the research support from the Institute of Economic Research (FIPE). The authors declare that they have no conflicts of interest.

[46] Finance Code 001

References

Angelucci, M. and D. Bennett: 2021, 'Adverse selection in the marriage market: HIV testing and marriage in rural Malawi'. *The Review of Economic Studies* **88**(5), 2119–2148.

Avert: 2018, 'HIV and AIDS in Brazil'. URL: https://www.avert.org/professionals/hiv-around-world/latin-america/brazil (accessed: 02/01/2021).

Bacon, O., M. L. Pecoraro, J. Galvão, and K. Page-Shafer: 2004, 'HIV/AIDS in Brazil'. *San Francisco: AIDS Policy Research Center, University of California*.

Barbosa, A. J., C. L. Szwarcwald, A. Pascom, and P. J. Souza: 2009, 'Trends in the AIDS epidemic in groups at highest risk in Brazil, 1980-2004'. *Cadernos de Saude Publica* **25**(4), 727–737.

Becker, G.: 1991, *A Treatise on the Family*. Harvard University Press, 2nd edition.

Brito, A. M. d., E. A. d. Castilho, and C. L. Szwarcwald: 2001, 'AIDS e infecção pelo HIV no Brasil: uma epidemia multifacetada'. *Revista da sociedade brasileira de medicina tropical* **34**(2), 207–217.

Brito et al., R.: 2015, 'The Brazilian response to HIV and AIDS'. URL: http://www.unaids.org/sites/default/files/country/documents/BRA_narrative_report_2015.pdf (accessed: 02/01/2021).

Bruze, G., M. Svarer, and Y. Weiss: 2015, 'The dynamics of marriage and divorce'. *Journal of Labor Economics* **33**(1), 123–170.

Chiappori, P.-A., B. Salanié, and Y. Weiss: 2017, 'Partner choice, investment in children, and the marital college premium'. *American Economic Review* **107**(8), 2109–67.

Choo, E.: 2015, 'Dynamic marriage matching: An empirical framework'. *Econometrica* **83**(4), 1373–1423.

Choo, E. and A. Siow: 2006, 'Who Marries Whom and Why'. *Journal of Political Economy* **114**(1), 175–201.

Cohen, D. J. and K. E. Bruce: 1997, 'Sex and mortality: Real risk and perceived vulnerability'. *Journal of Sex Research* **34**(3), 279–291.

Corno, L., N. Hildebrandt, and A. Voena: 2020, 'Age of marriage, weather shocks, and the direction of marriage payments'. *Econometrica* **88**(3), 879–915.

Darlington, S.: 2017, *Brazil Fights H.I.V. Spike in Youths With Free Preventive Drug.*

Dourado, I., S. MacCarthy, M. Reddy, G. Calazans, and S. Gruskin: 2015, 'Revisiting the use of condoms in Brazil'. *Revista Brasileira de Epidemiologia* **18**, 63–88.

Dourado, I., M. A. d. S. Veras, D. Barreira, and A. M. d. Brito: 2006, 'Tendências da epidemia de Aids no Brasil após a terapia anti-retroviral'. *Revista de Saúde Pública* **40**(suppl), 9–17.

Duflo, E., P. Dupas, and M. Kremer: 2006, *Education and HIV/AIDS Prevention: Evidence from a randomizedi evaluation in Western Kenya*, Vol. 4024. World Bank Publications.

Dupas, P.: 2011, 'Do Teenagers Respond to HIV Risk Information? Evidence from a Field Experiment in Kenya'. *American Economic Journal: Applied Economics* **3**(1), 1–34.

Fernandez, R., S. Parsa, and M. Viarengo: 2019, 'Coming out in America: AIDS, politics, and cultural change'. Technical report, National Bureau of Economic Research.

Francis, A. M.: 2008, 'The economics of sexuality: The effect of HIV/AIDS on homosexual behavior in the United States'. *Journal of Health Economics* **27**(3), 675–689.

Gao, X., Y. Wu, Y. Zhang, N. Zhang, J. Tang, J. Qiu, X. Lin, and Y. Du: 2012, 'Effectiveness of school-based education on HIV/AIDS knowledge, attitude, and behavior among secondary school students in Wuhan, China'.

Geoffard, P.-Y. and T. Philipson: 1997, 'Disease eradication: private versus public vaccination'. *The American Economic Review* **87**(1), 222–230.

Greenwood, J., P. Kircher, C. Santos, and M. Tertilt: 2017, 'The Role of Marriage in Fighting HIV: A Quantitative Illustration for Malawi'. *American Economic Review* **107**(5), 158–62.

Greenwood, J., P. Kircher, C. Santos, and M. Tertilt: 2019, 'An equilibrium model of the African HIV/AIDS epidemic'. *Econometrica* **87**(4), 1081–1113.

Grimard, F. and D. Parent: 2007, 'Education and smoking: Were Vietnam war draft avoiders also more likely to avoid smoking?'. *Journal of health economics* **26**(5), 896–926.

Guillon, M. and J. Thuilliez: 2015, 'HIV and Rational risky behaviors: a systematic review of published empirical literature (1990-2013)'.

Hankins, S. and M. Hoekstra: 2011, 'Lucky in life, unlucky in love? The effect of random income shocks on marriage and divorce'. *Journal of Human Resources* **46**(2), 403–426.

Hart, G. and L. Williamson: 2005, 'Increase in HIV sexual risk behaviour in homosexual men in Scotland, 1996–2002: prevention failure?'. *Sexually transmitted infections* **81**(5), 367–372.

Hauck, K.: 2018, 'The economics of infectious diseases'. In: *Oxford Research Encyclopedia of Economics and Finance*.

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

Laxminarayan, R. and A. Malani: 2006, 'Economics of infectious diseases'. In: *The Oxford handbook of health economics*.

Levi, G. C. and M. A. A. Vitória: 2002, 'Fighting against AIDS: the Brazilian experience'. *Aids* **16**(18), 2373–2383.

McFadden, D.: 1973, 'Conditional Logit Analysis of Qualitative Choice Behaviour'. In: P. Zarembka (ed.): *Frontiers in Econometrics*. New York, NY, USA: Academic Press New York, pp. 105–142.

Moatti, J.-P., N. Bajos, J.-P. Durbec, C. Menard, and C. Serrand: 1991, 'Determinants of condom use among French heterosexuals with multiple partners.'. *American Journal of Public Health* **81**(1), 106–109.

Murray, C. J., C. S. Callender, X. R. Kulikoff, V. Srinivasan, D. Abate, K. H. Abate, S. M. Abay, N. Abbasi, H. Abbastabar, J. Abdela, et al.: 2018, 'Population and fertility by age and sex for 195 countries and territories, 1950–2017: a systematic analysis for the Global Burden of Disease Study 2017'. *The Lancet* **392**(10159), 1995–2051.

Parker, R. and J. Galväo: 1996, 'Quebrando o silêncio: mulheres e AIDS no Brasil'. In: *Quebrando o silêncio: mulheres e AIDS no Brasil*.

Peltzman, S.: 1975, 'The effects of automobile safety regulation'. *Journal of political Economy* **83**(4), 677–725.

Posner, R. A. et al.: 1993, *Private choices and public health: The AIDS epidemic in an economic perspective*. Harvard University Press.

Rodríguez-Planas, N. and R. Tanaka: 2022, 'Gender norms and womenâs decision to work: evidence from Japan'. *Review of Economics of the Household* **20**(1), 15–36.

Schmidt, L.: 2008, 'Risk preferences and the timing of marriage and childbearing'. *Demography* **45**(2), 439–460.

Shapley, L. and M. Shubik: 1971, 'The assignment game I: The core'. *International Journal of Game Theory* **1**(1), 111–130.

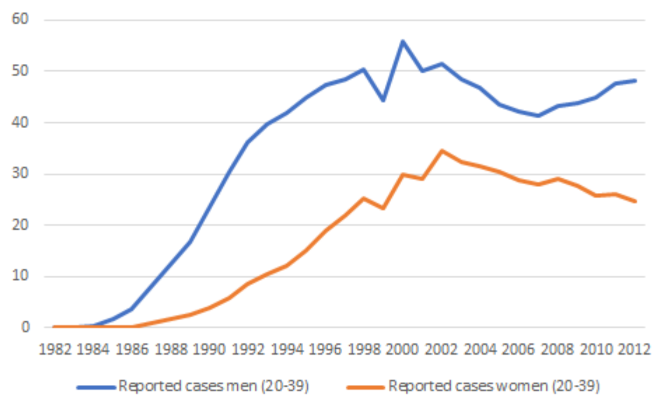
Spencer, M. K.: 2020, 'Safer Sex? The Effect of AIDS Risk on Birth Rates'. Mimeo.

Spivey, C.: 2010, 'Desperation or desire? The role of risk aversion in marriage'. *Economic Inquiry* **48**(2), 499–516.

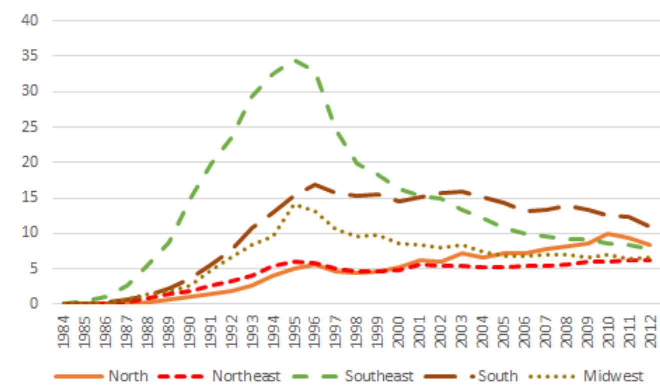
Szwarcwald, C. L., F. I. P. M. Bastos, E. A. d. Castilho, et al.: 1998, 'The dynamics of the AIDS epidemic in Brazil: a space-time analysis in the period 1987-1995'.

Tenkorang, E. Y., E. Maticka-Tyndale, and F. Rajulton: 2011, 'A multi-level analysis of risk perception, poverty and sexual risk-taking among young people in Cape Town, South Africa'. *Health & place* **17**(2), 525–535.

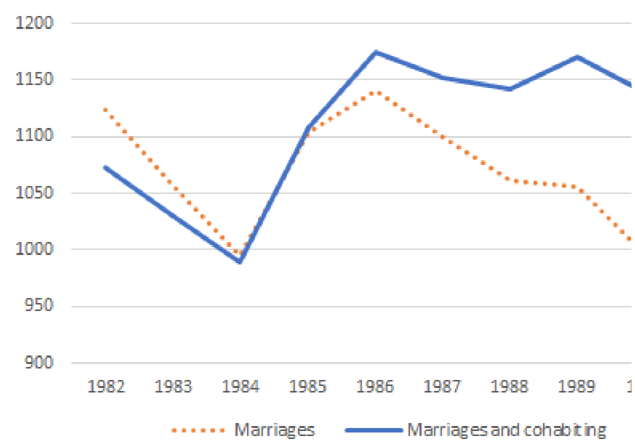
Tables and Figures



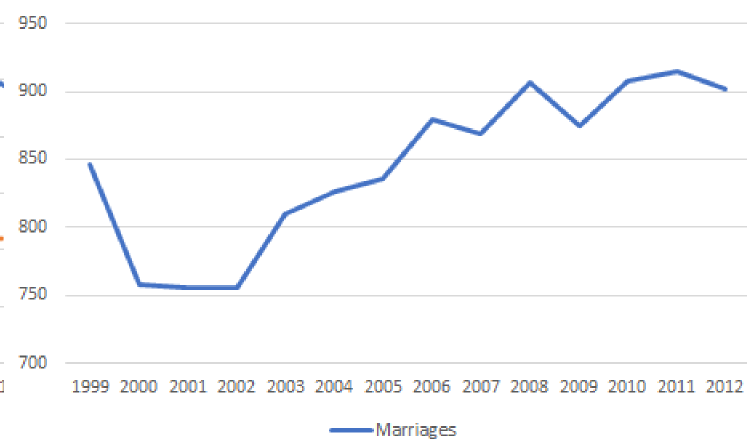
(a) Reported cases of HIV/AIDS by gender and ages of 20-39 years (per 100.000 inhabitants)



(b) HIV/AIDS deaths per region per 100.000 inhabitants (20-39 years)



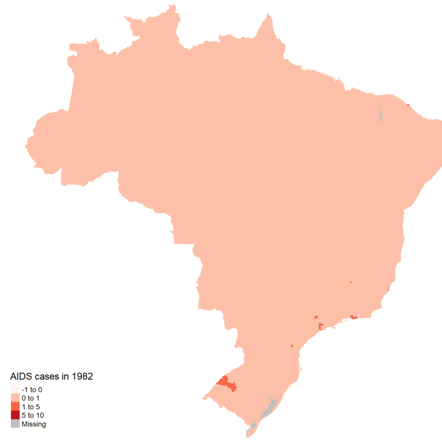
(c) Total marriages per 100.000 inhabitants (20-39 years)



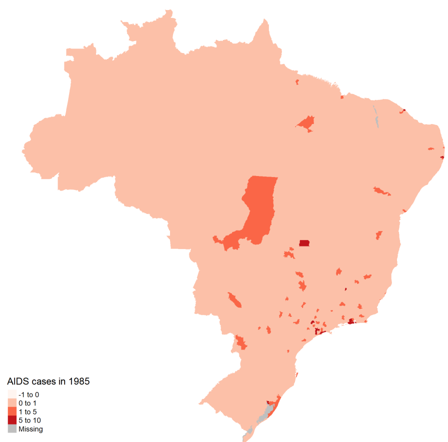
(d) Total marriages per 100.000 inhabitants (20-39 years)

Figure 1: Cases of HIV/AIDS and Marriages 100,000 inhabitants (20-39 years)

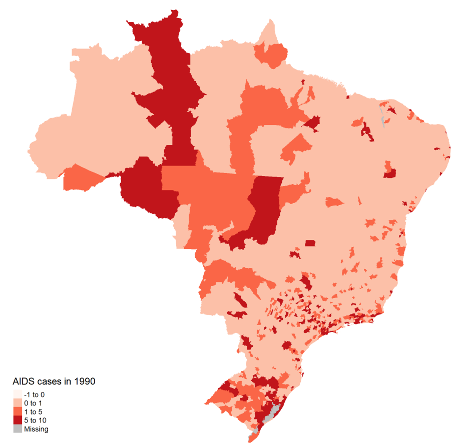
Sources: IBGE and Notification System of the Ministry of Health. In the figures, reported cases consider the year that the disease was diagnosed.



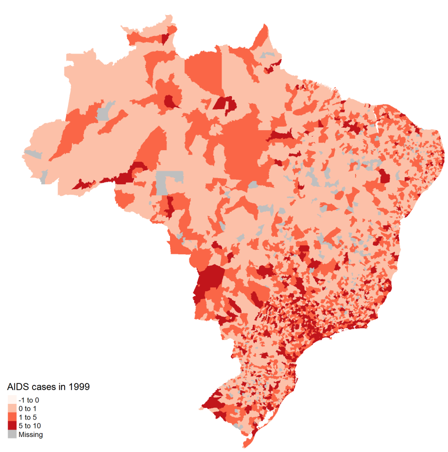
(a) Reported cases of HIV/AIDS by ages of 20-39 years and region in 1982



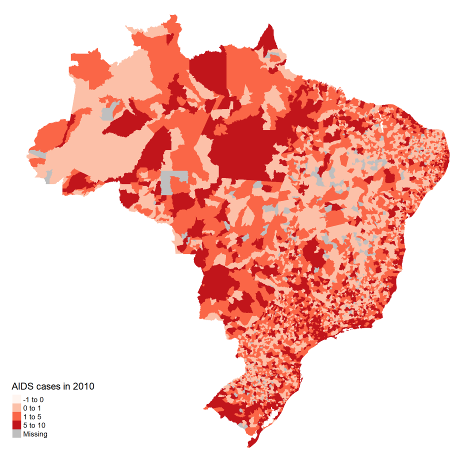
(b) Reported cases of HIV/AIDS by ages of 20-39 years and region in 1985



(c) Reported cases of HIV/AIDS by ages of 20-39 years and region in 1990



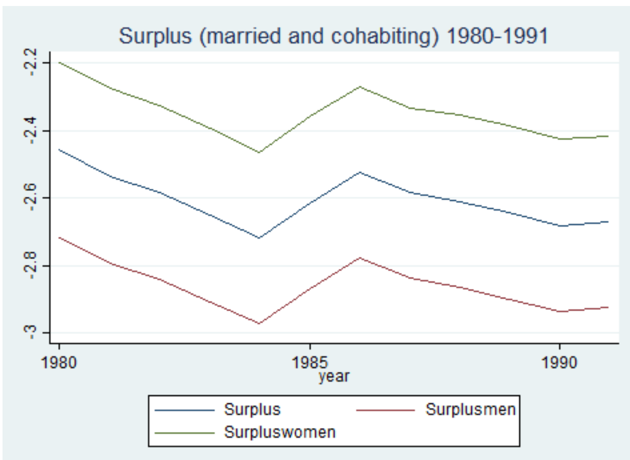
(d) Reported cases of HIV/AIDS by ages of 20-39 years and region in 1999



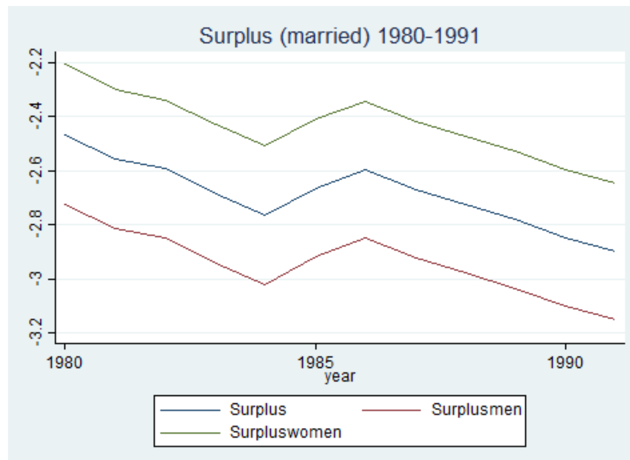
(e) Reported cases of HIV/AIDS by ages of 20-39 years and region in 2010

Figure 2: Reported cases of HIV/AIDS (20-39 years) by region

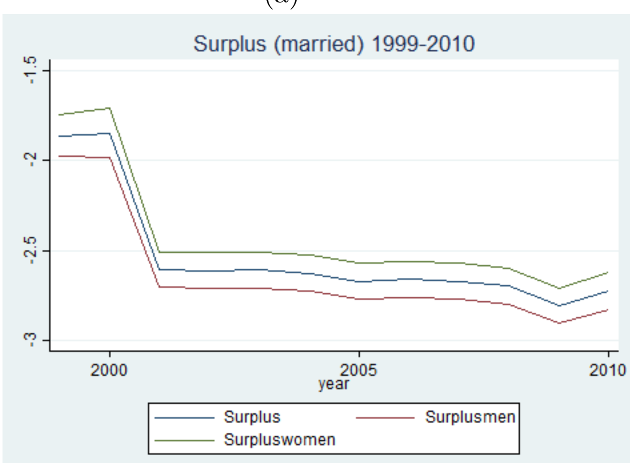
Sources: Notification System of the Ministry of Health. In the figures, reported cases consider the year that the disease was diagnosed.



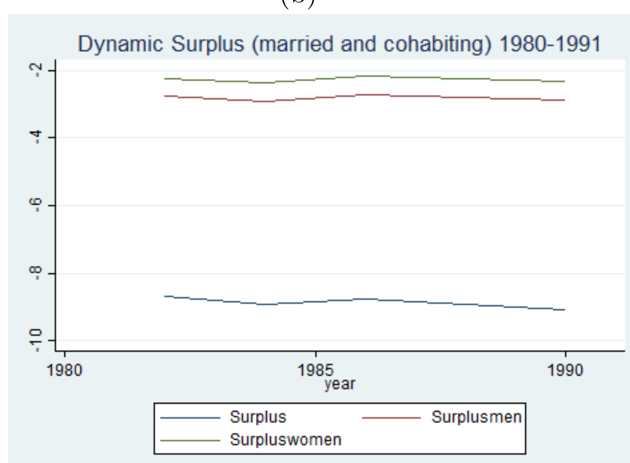
(a)



(b)



(c)



(d)

Figure 3: Couples' Surplus, Surplusmen and Surpluswomen from 20-39 years

Sources: IBGE and Brazilian Census 1991, 2000 and 2010.

Table 1: Impact of HIV/AIDS cases in elderly on cases in young adults, Phases 1 (1982-1991) and 2 (1999-2010)

	Cases of AIDS		
	1982-1991		1999-2010
	Married/Cohab. (1)	Married (2)	Married (3)
Cases of AIDS 65+ in t-1	1.14***	1.14***	0.02
	(0.44)	(0.44)	(0.01)
Cases of AIDS 65+ in t-2	1.78***	1.77***	0.03*
	(0.54)	(0.55)	(0.02)
Observations	35,075	33,810	35140
Kleibergen-Paap rk Wald F-statistics	37.07***	36.90 ***	6.11**
Mean Cases of AIDS	5.72	5.76	37.04

Notes: In Column (1) we consider first marriage and first cohabitation, in the next two columns we consider only first marriage. Columns (1) and (2) show results for Phase 1 (1982-1991), while Column (3) shows the results for Phase 2 (1999-2010). Cases of AIDS are the number of cases of AIDS per 100,000 inhabitants aged between 20 and 39 years old. The instruments are the lags of the cases of AIDS of people aged 65 or more in t-1 and in t-2. All regressions include fixed effects of year and municipality, mean age and years of schooling by municipality and state trends. Regressions are also weighted by total young adult population in the municipality. Standard errors (in parentheses) are clustered by municipalities.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Impacts of HIV/AIDS cases on marital surplus, OLS and IV estimates

	(1) OLS estimates	(2) IV estimates
Panel A: 1982-91 (married/cohab.)		
Cases of AIDS	0.075***	0.217***
	(0.03)	(0.08)
Observations	35,075	35,075
Kleibergen-Paap rk Wald F-statistics	-	37.07***
Hansen J statistic	-	1.17
Panel B: 1982-91 (married)		
Cases of AIDS	0.100***	0.353***
	(0.03)	(0.09)
Observations	33,810	33,810
Kleibergen-Paap rk Wald F-statistics	-	36.90***
Hansen J statistic	-	2.38
Panel C: 1999-2010 (married)		
Cases of AIDS	-0.019	-0.450
	(0.03)	(1.34)
Observations	35,140	35,140
Kleibergen-Paap rk Wald F-statistics	-	6.11**
Hansen J statistic	-	0.13

Notes: The dependent variable is the marital surplus (calculated in Equation 17). Panel A considers first marriage and first cohabitation, while Panels B and C consider only first marriage. Panels A and B show the results for Phase 1 (1982-1991) while Panel C shows results for Phase 2 (1999-2010). Cases of AIDS are the number of cases of HIV/AIDS per 100,000 inhabitants (young adults, 20-39 years). The instruments are the lags of cases of HIV/AIDS of people aged 65 or more in t-1 and t-2. Column (1) displays OLS estimates with controls (mean years of schooling, mean age, and state trends) and FEs (municipality and year fixed effects). Column (2) shows the IV estimates with municipality and year fixed effects and mean age by municipality and state-trend as controls. All regressions are weighted by total population in the municipality. Standard errors (in parentheses) are clustered by municipalities.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Impacts of HIV/AIDS cases on marital surplus by gender, IV estimates

	(1) Couple	(2) Women	(3) Men
Panel A: 1982-1991 (married/cohab.)			
Cases of AIDS	0.22***	0.25***	0.18**
	(0.08)	(0.08)	(0.08)
Observations	35,075	35,075	35,075
Kleibergen-Paap rk Wald F-statistics	37.07***	37.58***	35.91***
Hansen J statistic	1.17	1.25	1.14
Panel B: 1982-1991 (married)			
Cases of AIDS	0.35***	0.39***	0.32***
	(0.09)	(0.09)	(0.09)
Observations	33,810	33,810	33,810
Kleibergen-Paap rk Wald F-statistics	36.89***	37.39***	35.74***
Hansen J statistic	2.37	2.55	2.28
Panel C: 1999- 2010 (married)			
Cases of AIDS	-0.450	-0.434	-0.483
	(1.34)	(1.35)	(1.34)
Observations	35,140	35,140	35,140
Kleibergen-Paap rk Wald F-statistics	6.11**	6.14**	6.12**
Hansen J statistic	0.13	0.13	0.13

Notes: Panel A includes married and cohabiting couples. Panels B and C include only married couples. Panels A and B show the results for Phase 1 (1982-1991) while Panel C displays results for Phase 2 (1999-2010). Cases of AIDS are the number of cases of HIV/AIDS per 100,000 inhabitants (young adults from 20-39 years). The instruments are the lags in t-1 and t-2 of the cases of HIV/AIDS of people aged 65 or more. All regressions include fixed effects of year and municipality, mean age and years of schooling by municipality and state trend. Regressions are also weighted by the relevant population: column 1 by total population, and columns 2 and 3 by total population by gender in the municipality. Standard errors (in parentheses) are clustered by municipalities.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Impacts of HIV/AIDS incidence on marital surplus (married or cohabitation) by educational type and gender, IV estimates, 1982 to 1991

	Educational Type (Men/Women)			
	Low/Low	Low/High	High/Low	High/High
	(1)	(2)	(3)	(4)
Panel A: Couples				
Cases of AIDS	0.09	-0.02	-0.19	0.58***
	(0.11)	(0.23)	(0.29)	(0.16)
Kleibergen-Paap rk Wald F-stat.	37.53***	37.79***	37.27***	37.26***
Hansen J statistic	0.27	2.40	0.31	0.66
Observations	33,583	18,124	11,135	17,566
Panel B: Men				
Cases of AIDS	0.02	0.03	-0.41	0.45***
	(0.11)	(0.23)	(0.30)	(0.16)
Kleibergen-Paap rk Wald F-stat.	36.66***	37.06***	36.68***	36.50***
Hansen J statistic	0.32	2.36	0.34	0.76
Observations	33,583	18,124	11,135	17,566
Panel C: Women				
Cases of AIDS	0.17	-0.29	-0.14	0.70***
	(0.11)	(0.32)	(0.27)	(0.16)
Kleibergen-Paap rk Wald F-stat.	38.37***	38.67***	38.37***	38.00***
Hansen J statistic	0.22	0.35	1.63	0.55
Observations	33,583	18,124	11,135	17,566

Notes: Cases of Aids are the number of cases of aids per 100.000 inhabitants (young adults). The sample is restricted to data from Phase 1 (1982-1991). The instruments are the lags of the cases of AIDS of people aged 65 or more in t-1 and t-2. There are two levels of education: Low (elementary and middle school) and High (high school or above). Low/Low (men and women of type low); Low/High (men of type low and women of type high); High/Low (men of type high and women of type low); and High/High (when men and women are both of type high). All regressions include fixed effects of year and municipalities, mean age and mean years of schooling by municipalities. Regressions are also weighted by total population (Panel A), or total population by gender (Panels B and C) in the municipality. Standard errors are clustered by municipalities.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Impact of HIV/AIDS cases in elderly on cases in young adults, Phases 1 (1982-1991) and 2 (1999-2010), dynamic model

	Cases of AIDS		
	1982-1991		1999-2010
	Married and Cohab. (1)	Married (2)	Married (3)
Cases of AIDS 65+ t-2	2.86***	2.85***	0.011
	(1.08)	(1.11)	(0.02)
Observations	16,892	15,824	17,566
Kleibergen-Paap rk Wald F-statistics	9.82***	9.73***	0.26
Mean Cases of AIDS	5.61	5.67	36.60

Notes: In the dynamic model, the estimation is performed every two years. In the first column we consider first marriage and first cohabitation, in the next two columns we consider only first marriage. Columns (1) and (2) show results for Phase 1 (1982-1991), while Column (3) shows the results for Phase 2 (1999-2010). Cases of AIDS are the number of cases of AIDS per 100,000 inhabitants aged between 20 and 39 years old. The instrument is the lag of the cases of AIDS of people aged 65 in t-2. All regressions include fixed effects of year and municipality, mean age and years of schooling by municipality and state trend. Regressions are also weighted by total population in the municipality. Standard errors (in parentheses) are clustered by municipalities.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Impacts of HIV/AIDS cases on marital surplus by gender, Phase 1 (1982-1991) and 2 (1999-2010), dynamic model and IV estimates

	(1) Couples	(2) Women	(3) Men
Panel A: 1982-1991 (married/cohabitation)			
Cases of AIDS	0.16	0.13	0.06
	(0.12)	(0.12)	(0.13)
Observations	16,892	16,892	16,892
Kleibergen-Paap rk Wald F-statistics	9.82***	10.03***	9.60***
Panel B: 1982-1991 (married)			
Cases of AIDS	0.30**	0.30**	0.22
	(0.13)	(0.13)	(0.13)
Observations	15,824	15,824	15,824
Kleibergen-Paap rk Wald F-statistics	9.73***	10.00***	9.50***
Panel C: 1999-2010 (married)			
Cases of AIDS	0.04	0.10	-0.12
	(2.29)	(2.35)	(2.48)
Observations	17,566	17,566	17,566
Kleibergen-Paap rk Wald F-statistics	0.26	0.28	0.25

Standard errors in parentheses

Notes: The dependent variable is the dynamic marital surplus. In the dynamic model, the estimation is performed every two years following section 5.2. The dependent variables consider first marriage and first cohabitation. Cases of AIDS are the number of cases of HIV/AIDS per 100,000 inhabitants in every two years. The instrument is the lag cases of HIV/AIDS of people aged 65 in t-2. All columns displays IV estimates with controls and FEs, municipality and year fixed effects and mean age by municipality and state-trend as controls. All regressions are weighted by total population in the municipality. Standard errors (in parentheses) are clustered by municipalities..

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Impacts of HIV/AIDS incidence on marital surplus by educational type and gender, dynamic models and IV estimates, 1982 to 1991

	Types (Men/Women)			
	Low/Low	Low/High	High/Low	High/High
	(1)	(2)	(3)	(4)
Panel A: Couples (married/cohabiting)				
Cases of AIDS	-0.09	-0.13	-0.42	0.24
	(0.20)	(0.40)	(0.39)	(0.25)
Kleibergen-Paap rk Wald F-statistics	9.82***	9.88***	10.52***	9.90***
Panel B: Men				
Cases of AIDS	-0.27	-0.08	-0.57	0.29
	(0.21)	(0.41)	(0.38)	(0.25)
Kleibergen-Paap rk Wald F-statistics	9.82***	9.88***	10.52***	9.90***
Panel C: Women				
Cases of AIDS	-0.15	-0.15	-0.08	0.61***
	(0.19)	(0.40)	(0.38)	(0.23)
Kleibergen-Paap rk Wald F-statistics	9.82***	9.88***	10.52***	9.90***
Observations	16,187	8,268	4,818	8,097

Standard errors in parentheses

Notes: The dependent variable is the dynamic marital surplus. In the dynamic model, the estimation is performed every two years following section 5.2. The dependent variables consider first marriage and first cohabitation. Cases of Aids are the number of cases of AIDS per 100.000 inhabitants aged 20 to 39 years. The instrument is the lag of the cases of AIDS of people aged 65 or more in t-2. Regressions include fixed effects of year and municipalities, mean age and mean years of schooling by municipalities and state trends. All regressions are weighted by total population (in column 1) and total population by gender (in columns 2 and 3) in the municipality. Standard errors are clustered by municipalities.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Impacts of HIV/AIDS cases on marriage surplus using different ages as instruments

	Cases of AIDS		
	1982-1991	1982-1991	1999-2010
	Married and cohab.	Married	Married
	(1)	(2)	(3)
Panel A: IVs are lagged cases of AIDS 55+			
Cases of AIDS	0.182***	0.226***	1.413
	(0.05)	(0.06)	(5.11)
Kleibergen-Paap rk Wald F-statistics	53.79***	53.35***	0.39
Panel B: IVs are lagged cases of AIDS 60+			
Cases of AIDS	0.221***	0.290***	-0.991
	(0.05)	(0.06)	(1.62)
Kleibergen-Paap rk Wald F-statistics	42.33***	42.14***	4.86*
Panel C: IVs are lagged cases of AIDS 65+			
Cases of AIDS	0.217***	0.353***	-0.450
	(0.08)	(0.09)	(1.34)
Kleibergen-Paap rk Wald F-statistics	37.07***	36.88***	6.11**
Number of observations	35,075	33,810	35,140

Notes: Cases of AIDS are the number of cases of HIV/AIDS per 100,000 inhabitants aged between 20 and 39 years old. The instruments are the lags of the cases of AIDS of people aged 55 or more in t-1 and t-2 in Panel A, the lags in t-1 and t-2 with people aged 60 or more in Panel B and 65 or more in Panel C. All regressions include fixed effects of year and municipality, mean age and mean years of schooling by municipality and state trend. All regressions are weighted by total population in the municipality. Standard errors (in parentheses) are clustered by municipalities.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Impacts of HIV/AIDS cases on marriage surplus leaving one lagged instrument out

	(1) Marital Surplus	(2) Marital Surplus
Panel A: 1982-1991 (married)		
Cases of AIDS t-1	0.61***	
	(0.18)	
Cases of AIDS t-2		0.24*
		(0.13)
Kleibergen-Paap rk Wald F statistic	17.86***	22.30***
Observations	30,444	30,455
Panel B: 1999-2010 (married)		
Cases of AIDS t-1	0.15	
	(0.72)	
Cases of AIDS t-2		-0.25
		(0.63)
Kleibergen-Paap rk Wald F statistic	16.24***	24.15***
Observations	35,140	35,140

Notes: In Panel A and Panel B we consider first marriage. Cases of AIDS: number of cases of AIDS per 100,000 inhabitants aged between 20 and 39 years old in t-1, Column (1), and in t-2, Column (2). The dependent variable is marital surplus in t. Panel A presents the results from Phase 1 (1982-1991) while Panel B presents the results from Phase 2 (1999-2010). The regressions include fixed effects of year and municipality, mean age and years of schooling by municipality, state trend. All regressions are weighted by total population in the municipality. Standard errors are clustered by municipalities.

A Data Appendix

In this section, we describe in more detail the variables used in the empirical exercises. We use sampling weights to construct the variables.[47]

A.1 Marital Surplus

A.1.1 First Period 1982-1991: Marriage and Cohabitation

We keep in the 1991 Census base only people who declared themselves married (1st marriage or 1st cohabitation) or single and who married after 1977. The individuals who formed the base were aged between 20 and 39 years. For example, in 1985, only people born between 1946 and 1965 were included in the database. In addition, we create a variable "years of marriage" and keep in the sample the individuals who got married in the analyzed period or remained single. Moreover, we create a dummy if the person was married or single and then we sum the number of married and single people by gender, municipality and year, considering the sampling weight. Then we calculate the marriage surplus variables using equations 17, 18 and 19.

To estimate the dynamic equations we use the numerator of equations 20, 21 and 22 in t, every two years. We consider the married and cohabiting people. Moreover, we calculate the mean cases of aids in t and $t+1$ and we do the same to construct the instruments. Then, we calculate the denominator of the same equations to estimate the present value of being single in t and $t+1$. The final result is the present utility of being married in t compared to the present discounted sum of the payoffs from being single for t periods, where $t = 1, 2$.

A.1.2 First Period 1982-1991: Marriage

In this sample we consider only single people and those in their first marriage in order to be able to compare the results with the second period, where the data are available only for married people, that is, no information about cohabitation.

A.1.3 Second Period 1999-2010

During this period, we were unable to use the proportion of married people from the 2000 or 2010 Censuses because they no longer contained information about when the person married and if it was the first marriage (just whether people were married). Then we use the data available from the civil registries by age, gender, year and municipality, collected by IBGE.

[47] The microdata and dictionaries are available at https://www.ibge.gov.br/en/np-statistics/social/population/20620-summary-of-indicators-pnad2.html?=t=microdados. Microdata before 1999 are available at http://200.144.244.241:3004/. Additional information and dictionaries are available at http://www.econ.puc-rio.br/datazoom/english/censoMicro.html.

We keep in the 2000 and 2010 Census base only people who declared themselves single. The individuals who formed the base are aged between 20 and 39 years. For example, in 1999, only people born between 1960 and 1980 were included in the database. Moreover, we sum the proportion of single people by gender, municipality and year. In order to calculate the total singles in year *i* we add the married people in year *i+1*, gender and municipality. Hence, we have the total of single people that year.

Then we calculate the marital surplus variables using equations 17, 18 and 19.

A.2 Income

The dependent variable is the mean of income (tabulated in multiples of the minimum monthly waged, adjusted every year for inflation) in logarithm using data from the RAIS (Annual List of Social Information) aggregated by municipality and year from years 1985 to 1991 (the data was not available from 1982 to 1984). The age range used was 18 to 39 years. This range was the closest to what we use in our estimation (20-39). We obtained the number of inhabitants using Datasus data [48].

A.3 Levels of Education

We consider two levels of education: Low (elementary and middle school) and High (high school or above).

[48] http://www2.datasus.gov.br/DATASUS/index.php?area=02) to weight the regression by the number of Inhabitants

B Appendix

B.1 Model with complete information

In this section, we describe a two-period model with complete information (baseline for the incomplete information model) where men and women decide whether or not to date, and then if they will get married or remain single. If they remain single in the first period, they can enter in a new dating relationship in the second period and decide to get married or not. Men and women have symmetric preferences regarding marriage.[49]

There are two equally large populations of men and women to be matched. If people get married in the first period they will receive a payoff in both periods, but if they choose to marry in the second, they earn a payoff only in this period. The payoff of remaining single is zero. In this exercise (both sections), we do not include divorce or death rates similar to Angelucci and Bennett (2021).

The discount factor of a person i, δ_i, is idiosyncratic and distributed uniformly between 0 and 1: U $\sim$ [0,1].

The game follows these steps:

1. Man (woman) decides to enter or not in a dating relationship. Since we consider the payoff of entering a dating relationship is greater than remaining single, the agent decides to enter such a relationship. After that, nature plays and decides with probability p if the date is the right woman (man) and with probability (*1-p*) if the date is not the right person. The player observes the result and decides if he (she) will get married or not.

2. With probability p, he (she) will decide to get married and earn the payoff Π_r, where $\Pi_r > 0$. His (her) date observes his (her) decision and follows it, earning the same payoff.

3. With probability (*1-p*), he (she) will decide to get married with the wrong person or remain single and engage in another dating relationship. In the first case, he (she) will earn the payoff Π_w, where the payoff $\Pi_r > \Pi_w > 0$. In the latter case, he (she) decides to engage in another dating relationship and he (she) observes the result after nature plays again, where with probability p the new date is the right woman (man) or with probability (*1-p*) she (he) is not. In this last period he (she) remains single and earns zero or gets married and earns the payoff $p.\Pi_r + (1-p).\Pi_w$. As this expected payoff is positive and greater than zero, he (she) will always decide to get married instead remaining single.

Solving this game backwards, we have:

1. If nature chooses the date is the right person, than his(her) payoff is:

[49] This model does not include same sex marriage.

$$E(\Pi_i|t=1) = (1+\delta_i).[p.\Pi_r], \tag{23}$$

if this man (woman) decides to get married in the first period, or is

$$E(\Pi_i|t=2) = (\delta_i).p.[p.\Pi_r + (1-p).\Pi_w] \tag{24}$$

if he (she) decides to enter a new dating relationship and gets married only in the second period.

2. If nature chooses the date is the wrong one, than his (her) payoff is :

$$E(\Pi_i|t=1) = (1+\delta_i).[(1-p).\Pi_w], \tag{25}$$

if he (she) decides to get married in the first period.

Or, if he (she) enters in a new dating relationship, his (her) payoff is:

$$E(\Pi_i|t=2) = (\delta_i).(1-p).[p.\Pi_r + (1-p).\Pi_w] \tag{26}$$

B.2 The decision on marriage

If nature chooses the date is the right person, which is observed by the agent, he (she) decides to get married in the second period if

$$E(\Pi_i|t=2) = (\delta_i).p.[p.\Pi_r + (1-p).\Pi_w] > (1+\delta_i).[p.\Pi_r]. \tag{27}$$

Therefore, for any $\delta_i > 0$ and from equation (27), the dominant strategy is for the agent to always marry in the first period as $\Pi_r > p.\Pi_r + (1-p).\Pi_w$.

In the case where nature chooses the date is the wrong one, man (woman) will get married in the second period only if:

$$(\delta_i).(1-p).[p.\Pi_r + (1-p).\Pi_w] > (1+\delta_i).[(1-p).\Pi_w] \tag{28}$$

That is, $\delta_i > \Pi_w/(p[\Pi_r - \Pi_w])$, otherwise he (she) gets married in the first period.

C Appendix

C.1 The decision on marriage

A High-type person marries in the first period if the expected payoff of people not infected by the virus, when nature chooses the date is the right person in the first period, is greater than if marrying in the second. He (she) will get married if:

$$E(\Pi_i^h|t=1) = (1+\delta_i).p.[\Pi_r^{hh}.(1-\phi_1(X)) + \Pi_r^{hl}.\phi_1(X)] > \\ p.\delta_i.\{[p.[(1-\phi_2(X)).\Pi_r^{hh} + \phi_2(X).\Pi_r^{hl})] + (1-p).[(1-\phi_2(X)).\Pi_w^{hh} + \phi_2(X).\Pi_w^{hl})]\} \tag{29}$$

A Low-type person marries in the first period if the expected payoff of people infected by the virus, when nature chooses the date is the right person in the first period, is greater than if marrying in the second. He (she) will get married if:

$$E(\Pi_i^h|t=1) = (1+\delta_i).p.[\Pi_r^{lh}.(1-\phi_1(X)) + \Pi_r^{ll}.\phi_1(X)] > \\ p.\delta_i.\{[p.[(1-\phi_2(X)).\Pi_r^{lh} + \phi_2(X).\Pi_r^{ll})] + (1-p).[(1-\phi_2(X)).\Pi_w^{lh}] + \phi_2(X).\Pi_w^{ll})]\} \tag{30}$$

Agents will decide to marry in the first or second period depending on the discount factor δ_i. If the threshold is as in equation (31), they will marry in the first period:

$$\delta_i < [\Pi_r^{hh}.(1-\phi_1(X)) + \Pi_r^{hl}.\phi_1(X)]/\{p.[(1-\phi_2(X)).\Pi_r^{hh}] + \\ \phi_2(X).\Pi_r^{hl}] + (1-p)[(1-\phi_2(X)).\Pi_w^{hh}] + \phi_2(X).\Pi_w^{hl}] - \\ [(1-\phi_1(X)).\Pi_w^{hh} + \phi_1(X).\Pi_w^{hl}]\} \tag{31}$$

A High-type person marries in the second period if the expected payoff of people not infected by the virus, when nature chooses the date is the wrong person in the first period, is greater than of marrying in the first. He (she) will enter in a new dating relationship and get married if:

$$(1-p).\delta_i.\{p.[(1-\phi_2(X)).\Pi_r^{hh} + \phi_2(X).\Pi_r^{hl})] + (1-p).[(1-\phi_2(X)).\Pi_w^{hh} + \phi_2(X).\Pi_w^{hl})\} > \\ (1-p).(1+\delta_i).[(1-\phi_1(X)).\Pi_w^{hh} + \phi_1(X).\Pi_w^{hl})] \tag{32}$$

In this case the agent will decide to marry in the first period when the threshold δ_i is:

$$\delta_i > 1/\{p.[(1-\phi_2(X)).\Pi_r^{hh}] + \phi_2(X).\Pi_r^{hl}] + (1-p)[(1-\phi_2(X)).^{hh}] + \phi_2(X).\Pi_w^{hl}] - \\ [(1-\phi_1(X)).^{hh}] + \phi_1(X).\Pi_w^{hl}]/[(1-\phi_1(X)).^{hh}] + \phi_1(X).\Pi_w^{hl}]\} \tag{33}$$

A Low-type person marries in the second period if the expected payoff of people not infected by the virus, when nature chooses the date is the wrong person in the first period, is greater than of marrying in the first. He (she) will enter in a new dating relationship and get married if:

$$(1-p).\delta_i.\{p.[(1-\phi_2(X)).\Pi_r^{lh} + \phi_2(X).\Pi_r^{ll})] + (1-p).[(1-\phi_2(X)).\Pi_w^{lh}] + \phi_2(X).\Pi_w^{ll})\} > \\ (1-p).(1+\delta_i).[(1-\phi_1(X)).\Pi_w^{lh} + \phi_1(X).\Pi_w^{ll})] \tag{34}$$

The decision of getting married in the first or second period may change when the probability of dating someone infected by AIDS changes. This we show in the next section.

D Appendix Tables

Appendix Table 1: Impacts of HIV/AIDS cases of elderly (+55, +60 and +65 years) on cases in young adults, Phase 1 and 2

	(1) Cases of AIDS 1982-91	(2) Cases of AIDS 1999-2010
Panel A: 55+		
Cases of AIDS 55+ t-1	1.54***	-0.00
	(0.25)	(0.14)
Cases of AIDS 55+ t-2	1.79***	0.09
	(0.33)	(0.16)
Kleibergen-Paap rk Wald F statistic	23.43 ***	0.20
Observations	35,075	35,140
Mean of dep. var.	5.72	37.04
Panel B: 60+		
Cases of AIDS 60+ t-1	1.65***	0.02
	(0.36)	(0.01)
Cases of AIDS 60+ t-2	2.15***	0.03
	(0.42)	(0.02)
Kleibergen-Paap rk Wald F statistic	16.06 ***	1.97
Observations	35,075	35,140
Mean of dep. var.	5.72	37.04
Panel C: 65+		
Cases of AIDS 65+ t-1	1.14***	0.02
	(0.44)	(0.01)
Cases of AIDS 65+ t-2	1.78***	0.03*
	(0.54)	(0.02)
Kleibergen-Paap rk Wald F statistic	5.98***	1.97
Observations	35,075	35,140
Mean of dep. var.	5.72	37.01

Notes: Column (1) considers first marriage and cohabitation and Column (2) only first marriage due to data restrictions. Cases of AIDS are the number of cases of HIV/AIDS per 100,000 inhabitants aged between 20 and 39 years old. We use different age ranges for the instruments: lags of the cases of AIDS of people aged 55 or more, 60 or more, and 65 or more in t-1 and t-2. All regressions include fixed effects of year and municipality, mean age by municipality and state trend. All regressions are weighted by total population in the municipality. Standard errors (in parentheses) are clustered by municipalities. Kleibergen-Paap rk Wald F-statistics report the first-stage test of the excluded instruments.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix Table 2: Impacts of HIV/AIDS cases on hospital beds, IV estimates, 1982-1991.

	(1) Hospital Beds
Panel A: 1982-90 (married and cohab), static model	
Cases of AIDS	-4.38
	(4.22)
Observations	28,010
Kleibergen-Paap rk Wald F-statistics	24.19***
Panel B: 1982-90 (married and cohab), dynamic model	
Cases of AIDS	-2.89
	(2.83)
Observations	13,427
Kleibergen-Paap rk Wald F-statistics	11.97***

Notes: Dependent variable is the number of hospital beds per 100.000 inhabitants in the municipality. Cases of AIDS represents the number of cases of HIV/AIDS per 100.000 inhabitants. The regression includes mean age and mean years of schooling by municipality, municipality and year fixed effects and state-specific trend. The static regression includes the following instrumental variables: the lags of the cases of AIDS of people aged 65 or more in t-1 and t-2. The dynamic regression includes the following instrumental variable: the lag of the cases of AIDS of people aged 65 or more in t-2. All regressions are weighted by total population in the municipality. Standard errors are clustered by municipalities.

Appendix Table 3: Impacts of HIV/AIDS cases on Income (total and by gender), IV estimates, 1985-1991

	(1) Average income	(2) Income (women)	(3) Income (men)
Cases of AIDS	-0.000	-0.001	-0.000
	(0.00)	(0.00)	(0.00)
Observations	24,467	24,025	24,404
Kleibergen-Paap rk Wald F-statistics	38.28 ***	38.59 ***	37.15 ***
Hansen J statistic	0.21	0.54	0.83

Notes: The dependent variables of Columns (1) to (3) are mean of income (in multiples of the minimum monthly wage) in logarithm for all samples (Column 1), mean of income for women (Column 2) and mean of income for men (Column 3). Cases of AIDS represent the number of cases of HIV/AIDS per 100,000 inhabitants. All regressions include mean age and mean years of schooling by municipality, municipality and year fixed effects and state-specific trend. We instrument cases of HIV/AIDS for young adults using the lags of the cases of AIDS of people aged 65 or more and in t-1 and t-2. We also present the F-statistic of the first-stage results. All regressions are weighted by total population in the municipality. Standard errors (in parentheses) are clustered by municipalities.

Appendix Table 4: Impacts of HIV/AIDS cases on marital surplus with and without income as control variable, IV estimates, 1982-1991

	(1) Marital Surplus	(2) Marital Surplus
Cases of AIDS	0.22***	0.32**
	(0.078)	(0.133)
Income		-0.03
		(1.309)
Observations	35,075	24,467
Kleibergen-Paap rk Wald F-statistics	38.60 ***	37.15***

Notes: The dependent variable is the marital surplus for the entire sample - first marriage and cohabitation (Column 1 and 2). The difference between columns 1 and 2 is that in the second we add average income in logarithm as an additional control, which is not statistically significant. Cases of AIDS represent the number of cases of HIV/AIDS per 100,000 inhabitants. The instruments are the lags of the cases of AIDS of people aged 65 or more in t-1 and t-2. We also present the F-statistic of the first-stage results. All regressions include year and municipality fixed-effects, state-specific trends, and controls (mean age by municipality). All regressions are weighted by total female population in the municipality. Standard errors (in parentheses) are clustered by municipalities.

Appendix Table 5: First stage: Impacts of HIV/AIDS cases of elderly (+65 years) on cases in young adults, Phase 2

	(1) Cases of AIDS 1999-2010	(2) Cases of AIDS 1999-2010
Cases of AIDS 65+ t-1	0.02	0.03*
	(0.01)	(0.02)
Cases of AIDS 65+ t-2	0.03*	0.04**
	(0.02)	(0.02)
Cases of AIDS 65+ t		0.06***
		(0.02)
Observations	35,140	35,140
F-stat	6.11**	24.6***
Hansen J statistic	0.13	0.13

This table includes only married couples and displays results for Phase 2 (1999-2010). Cases of AIDS are the number of cases of HIV/AIDS per 100,000 inhabitants (young adults from 20-39 years). The instruments are the lags in t-1 and t-2 of the cases of HIV/AIDS of people aged 65 or more (column 1) and in column 2 we add the contemporary cases of elderly. All regressions include fixed effects of year and municipality, mean age and years of schooling by municipality and state trend. Regressions are also weighted by the relevant population: total population in the municipality. Standard errors (in parentheses) are clustered by municipalities.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$