The Effect of HIV Infection Risk Beliefs on Risky Sexual Behavior: Scared Straight or Scared to Death?

Jason T. Kerwin[1]
University of Michigan Department of Economics and Population Studies Center
April 2014

Economists typically assume that risk compensation has uniformly negative or self-protective effects – that people become more careful as the perceived health risks of their actions increase. However, under certain conditions, rational responses can instead be positive, or fatalistic: increased perceived risks can lead people to take *fewer* precautions. By extending the basic model of rational responses to health risks, I show that fatalism will result whenever perceived health risks are sufficiently high and people do not have perfect control over all possible exposures. This result holds even for agents who do not understand probability in a sophisticated way. I then test for the possibility that sufficiently high health risks induce fatalistic responses using a randomized experiment that provided information on HIV transmission risks to people in rural Malawi. On average, responses to the information are self-protective and fairly small in magnitude, but statistically significant: the risk elasticity of sexual behavior is roughly -0.5. Consistent with my theoretical predictions, these responses are strongly heterogeneous by baseline risk beliefs. I develop a novel estimation strategy that extends the logic of heterogeneous treatment effects analysis to instrumental-variables estimates of the marginal effect of risk beliefs on sexual risk-taking. Using this approach, I find that over 10% of the population has risk beliefs high enough that they have become fatalistic, so their risk elasticity of sexual behavior is positive rather than negative. These individuals reduced their sexual risk-taking in response to the information treatment, consistent with the notion that HIV risk messaging is backfiring for this subpopulation. They also have higher baseline indicators of sexual activity, which suggests that the consequences of providing accurate risk information for the overall prevalence of HIV are epidemiologically ambiguous.

[1]I am grateful to Rebecca Thornton, Jeff Smith, John DiNardo, David Lam, Susan Watkins, Ophira Vishkin, Evan Herrnstadt, Joe Golden, Audrey Dorélien, Laura Derksen, and seminar participants at the University of Michigan, MEA, and PacDev for their valuable feedback. Data collection on this project was supported by the Center for Global Health, the Population Studies Center, the Weinberg International Travel Fund, the Center for Education of Women, and the Michigan Institute for Teaching and Research in Economics at the University of Michigan as well as the Russell Sage Foundation's Small Grants in Behavioral Economics program. I acknowledge fellowship support from the Population Studies Center and the Rackham Graduate School. All remaining mistakes are my own.

1. Introduction

Risk compensation is central to our understanding of how people make decisions about protecting their health. Decisions about many activities, ranging from smoking and drinking to risky sex to merely driving to work, involve comparing the definite benefits of an act against uncertain, but potentially large, negative consequences. Beginning with the seminal Peltzman (1975) paper on automobile safety regulation, economists have realized that declines in the risks associated with certain behaviors are often offset by rational increases in risk-taking. Empirical research on risk compensation typically assumes that the sign of the effect is uniformly negative – that when the per-act risk goes up, people take fewer chances, a pattern that can be described as self-protective. This paper considers the possibility that for some people rational responses to health risks are instead "fatalistic" – that the optimal choice may be to *increase* one's risk-taking when per-act risks rise.

Most empirical research on health behaviors assumes that risk compensation is always self-protective, and ignores the possibility that some people's responses could have the opposite sign. A prediction of uniformly self-protective risk compensation can be derived by assuming agents optimize their level of risk-taking given a fixed per-act risk of a bad outcome (such as an HIV infection), with the simplifying assumption that individual risks are linearly additive.[2] Based on this model, regressions of health behaviors on risks typically assume that the risk term enters linearly into the regression function.[3] This hides any potential non-monotonicity in the relationship being studied – if the relationship is in fact self-protective for most people, but fatalistic for some of them, what results is simply a negative coefficient estimate with a somewhat lower magnitude.

The common assumption by empirical researchers that risk compensation is monotonically negative contrasts sharply with a theoretical tradition that shows that for certain people, rational responses to increases will instead be fatalistic.[4] The fundamental idea in all of

[2] This model is explicit in Oster (2012), but is used implicitly in many empirical analyses which restrict the relationship between risks and behavior to be linear and therefore monotonic.

[3] An extreme example is the Viscusi (1990) study of cancer risk perceptions and smoking behavior, which employs one-sided rather than two-sided t-tests. This eliminates any possibility of fatalistic responses, although the estimated standard errors are small enough that this assumption does not affect inference.

[4] Kremer (1996) was the first to develop this idea: he shows that if we relax the assumption that the expected cost of HIV infection is linear, the effect of HIV prevalence on the marginal cost of having more sex partners is non-monotone. O'Donoghue and Rabin (2001) show that the binomial CDF has a cross-partial derivative with respect to

these models is that a change in the per-act risk affects not only the marginal cost of the acts the agent is deciding over, but also a stock of previously-chosen acts over which one no longer has any control. If an agent's perceived per-sex-act risk of contracting HIV rises, this has a direct effect of increasing the marginal cost (in expected utility) of having more risky sex. But it also increases the probability that the agent *already* has HIV, which *decreases* the marginal cost of more risky sex. When the second effect dominates, increases in perceived risks will lead to more risk-taking rather than less. This idea can be extended to include future acts which one cannot possibly avoid. For example, if a person is using condoms to avoid HIV infection and understands that they sometimes break, then a sufficiently high risk of infection from unprotected sex will mean that you are doomed no matter what you do.

Section 2 of this paper illustrates this insight through a simple extension to the typical model employed in most empirical work: rather than using a linear approximation to the total probability of a bad outcome, I instead employ general function that satisfies the basic properties of valid probabilities. This function admits the true function, based on the binomial CDF, as a special case. I use this model to formalize the insight described above. I also extend the existing theoretical work on fatalism in two crucial ways. First, I show that fatalistic risk compensation does not rely on people understanding how individual risks actually combine to form the total probability of an HIV infection. Rather, it holds whenever people understand that probabilities are bounded above by one. Second, I prove that fatalism can hold for interior solutions, rather than just corner cases where people choose the maximal number of risky acts.

My theoretical framework suggests that people with sufficiently high beliefs about the risks of their behaviors, and imperfect control of their entire risk-taking history, will tend to become fatalistic, because they feel that they are doomed irrespective of what they do. One health behavior where these conditions are plausible is risky sex, in a context where HIV is a major problem. In particular, there is suggestive evidence that some people are behaving fatalistically with respect to the risk of HIV infection in Malawi, a country in the Southern Africa that has the ninth-highest HIV prevalence in the world.[5] Official HIV prevention messaging in Malawi, as in most countries, encourages people to hold extremely high estimates of

---

the number of sex acts and the per-act risk that is not uniformly positive, which drives fatalistic responses for people with sufficiently high beliefs.

[5] Estimates of Malawi's HIV prevalence vary depending on the source. Primary data from the 2010 Demographic and Health Survey shows a prevalence of 10.6% among adults aged 18-49.

transmission rates. Figure 1 excerpts a "case study" from a textbook used in Malawi to teach HIV prevention (yellow highlighting added); this case study is the only information about transmission rates that I was able to find in official secondary school textbooks. It suggests the transmission rate is 100%, which is much higher than the true risk of about 0.1% per unprotected sex act.[6] As a result of this messaging, people in Malawi greatly overestimate both the prevalence and transmission rate of HIV: the median person in my sample thinks 55% of their neighbors have HIV, and that the transmission rate of the virus from a single unprotected sex act is 100%. Despite these high risk beliefs, people in rural Malawi have unprotected sex nearly 90% of the time, suggesting that net risk compensation is extremely limited. One possible explanation for this apparently limited response is that some people are responding fatalistically, responding to the high risks by being less careful instead of more so. This is consistent with qualitative research in Malawi, which has found that some men in the country employ fatalistic reasoning when talking with each other about risky sex (Kaler 2003).[7]

To test the possibility of non-monotone responses to HIV risks, I utilize a field experiment that I conducted in a rural area of Southern Malawi's Zomba District in 2012. The experiment recruited 1292 respondents from 70 villages, and randomly assigned 35 of the villages to be taught medically-accurate information about HIV transmission risks. A baseline survey was conducted prior to the information treatment, and a followup survey was conducted four to twelve weeks later. The experiment is described in detail in Section 3, which also discusses my choice about which outcome measures to use. I rely primarily on self-reports of sexual behavior from a retrospective sex "diary" that is designed to reduce recall bias and mitigate social acceptability bias.

In Section 4, I turn to the results of the experiment. The randomized information treatment decreased people's beliefs about the risks of unprotected sex substantially: at the endline survey, the average person in the treatment group believed the risk of HIV transmission

[6] Aside from the textbooks, other school materials I have collected from Malawian classrooms, such as comic book dramas on the dangers of AIDS, also imply a 100% transmission rate. Despite the fact that most Malawians never reach secondary school, it appears that the information promulgated by the Life Skills course does spread to the population as a whole. For instance, the aforementioned case study implies that life expectancy from the moment one contracts HIV is 3 years (as opposed to the true figure of roughly 10), which is quite close to what the average Malawian believes.

[7] This is in line with is the selfishly rational, fatalistic-style reasoning suggested by the passage: after learning she is HIV-positive, Nabetha "decided she would not suffer alone but infect as many people as possible. She worked very hard toward this goal." While the intent of the case study is clearly to have students discuss why what Nabetha did was wrong, the passage inadvertently models fatalistic reasoning.

was 33% per sex act, as opposed to 74% in the control group. The experiment also allows me to isolate the causal effect of the information treatment on people's sexual behavior. The average treatment effect was small, but statistically significant. I find a point estimate of slightly less than 0.1 SDs of the control-group level of sexual activity, and can rule out effects bigger than 0.2 SDs. Using the experimental treatment as an instrumental variable, I estimate the effect of people's risk beliefs on risky sex. I find that these effects are also small, but statistically different from zero: the average risk elasticity of risk-taking is about -0.5.

To take the data to the model, I begin with simple heterogeneous treatment effects tests. These reveal substantial heterogeneity in the responses to the information treatment by people's baseline risk beliefs: people with initially low risk beliefs respond positively to the new information (which lowers their risk beliefs), while people with initially high risk beliefs respond *negatively*. This is the same non-monotonic pattern of responses predicted by the model of rational fatalism. Moreover, there is no evidence of significant heterogeneity in treatment effects by any other plausible baseline factor, such as gender or previous sexual activity.

If responses to risks are rationally fatalistic, this should result in heterogeneity by baseline risk beliefs not only in the effect of the information treatment, but also in the marginal effect of risk beliefs on risky behavior. To explore this I develop a method for decomposing 2SLS estimates of marginal effects by exogenous covariates. I show that this method gives consistent estimates of the underlying conditional parameter, and then apply it using both a partially non-linear regression and a bracketed linear regression. Both sets of results suggest significantly negative risk responses for people with low risk beliefs, and significantly positive risk responses for people with high risk beliefs. The people who exhibit positive, fatalistic responses also have the highest baseline number of lifetime sex partners, suggesting that they may be more important in driving the overall prevalence of HIV. This means that the effect of the status quo policy – in which health educators encourage people to greatly overestimate HIV transmission risks, for their own good – is ambiguous from both an ethical and an epidemiological standpoint. More generally, these results mitigate against the programs that attempt to "scare people straight" via messages that emphasize that risks are extremely high – especially when they actually are not.

In Section 5.1, I examine the mechanisms behind the fatalism observed for people with initially high risk beliefs. I show that there is no evidence that fatalistic responses were driven by

people's beliefs about their current HIV status – which is entirely plausible, given the widespread availability of HIV testing in the local area. In contrast, I do find evidence that the same group of people who reduced their level of risk-taking in response to the information treatment also became less likely to believe they were doomed to contract HIV in the future. This suggests that in this population, fatalism was operating through imperfect control over future risk-taking, rather than past risky acts that one has already committed. Section 5.2 discusses the implications of these results for models of rational epidemics. Research in economic epidemiology has demonstrated that risk compensation plays a key role in shaping how certain diseases spread. This literature has focused in particular on HIV because it is spread almost entirely through voluntary behavior (Philipson and Posner 1993). Meanwhile, traditional HIV epidemiology has shown that epidemics of the virus are driven by high-activity groups. My results from Section 4 show that the self-protective pattern of risk compensation assumed by many studies of HIV and sexual behavior is reversed among a group of people with high risk beliefs. I also show that the fatalistic group has higher baseline indicators of risky sexual behavior than the rest of the population. This suggests that they are likely to be over-represented in the high-activity groups that drive HIV epidemics. As a result, the epidemiological implications of my information treatment are ambiguous: it is possible that an information campaign revealing the true risk of HIV infection could reduce over HIV prevalence, because of its effects on the fatalistic group. In Section 5.3 I discuss the experiment's direct effect on HIV infections, which was sufficiently small as to be offset by the subsidized condoms provided during the endline survey.

In Section 5.4 I discuss the limitations of this study. The most important potential limitation – which is that my heterogeneous treatment effects analyses could constitute atheoretical "fishing" for results – is mitigated by the fact that the motivating theory behind this paper was derived prior to collecting the data for this project, along with key stylized facts about the HIV epidemic in Malawi (Kerwin 2012 WP).

This paper contributes to four bodies of research in economics. First, it builds on our understanding of risk compensation by providing what I believe to be the first experimental evidence on the elasticity of risky behavior with respect to perceived risks. Moreover, it shows that that elasticity cannot be meaningfully summarized by a population average, because the subgroup of the population with the highest baseline risk beliefs responds positively

(fatalistically) to risks. This implies that future empirical work on risk compensation should take into account the possibility of fatalism.

Second, it contributes to a growing empirical literature that studies how people's subjective expectations affect their behavior. Expectations have long played an important role in economic models, but efforts to measure what people actually believe are relatively new. Recent research has shown that it is possible to collect meaningful information on people's subjective expectations can both in the developed world[8] as well as in developing countries.[9] I demonstrate that these subjective beliefs have a measurable, causal effect on people's behavior, lending credence to the broader idea that we should be asking people about their subjective beliefs rather than assuming they know the true probabilities of events.

Third, it relates to research on the rational models of HIV testing and sexual behavior. As Philipson and Posner (1993) point out, the effect of learning one's HIV status is theoretically ambiguous, because learning that you are HIV-positive can have two opposite-signed effects on your behavior. Purely self-interested people should see little or no marginal cost from further risky sex if they are already infected, while altruistic people would want to take measures to protect their prospective partners. A parallel logic applies to those who learn they are HIV-negative. Experimental research on HIV testing and sexual behavior has found conflicting results: Thornton (2008) finds that effects on sexual behavior are only statistically distinguishable from zero for people who learn they are HIV-positive: HIV-positives reduce their sexual risk-taking, but by a very small amount. In contrast, Gong (2013) finds that people who are surprised by an HIV-positive result increase their level of risky sexual behavior as a result. While my experiment does not directly study HIV testing, the finding that some people are fatalistic may help inform how we think about these results. In particular, both results are reconcilable with a model in which HIV-positive people are altruistic, but some are fatalistic about transmitting the virus to their sex partners.

Fourth, it helps explain an empirical mystery in the study of risk compensation and HIV. Studies of the US HIV epidemic have consistently shown important responses to the virus; for example, that gay men in the United States reduced their sexual risk-taking sharply as the prevalence of HIV rose during the 1980s. This pattern has been much less evident in Africa,

[8] Lillard and Willis (2001)
[9] Attanasio (2009)

which is home to the vast majority of HIV-infected people.[10] The differential pattern in Africa could be explained in part by fatalism, due to messaging that emphasizes extremely high risks. However, two recent randomized trials have found that certain groups of people in Africa do respond to information about relative HIV transmission risks by adapting their sexual behavior. Godlonton et al. (2012) find that uncircumcised men in Malawi take fewer sexual risks when they are told that circumcised men face a lower risk of HIV infection. In a study in Kenya, Dupas (2011) finds that girls in secondary school choose younger partners when they are told that older partners are riskier. Rational fatalism suggests an alternative explanation that could account for both the gap in responses between the US and Africa, and the difference between the limited population-average response and the strong responses from some population groups. If a negative, self-protective response by some is balanced by a positive, fatalistic response by others, the average response will be lower in magnitude, and potentially near zero, unless the researcher explicitly allows for non-monotonic responses. Likewise, fatalism by other groups could offset strong responses by the groups studied in the two experiments described above to generate the small overall responses seen across Africa.

The remainder of this paper is organized as follows: I begin in Section 2 by laying out a model of responses to risks that extends the typical approach to allow for the possibility of rational fatalism, showing that under very general conditions people may rationally respond to high perceived risks fatalistically (as opposed to self-protectively). In Section 3, I describe a randomized field experiment that I conducted in Southern Malawi to test the implications of this model, as well as the data on risk beliefs and sexual risk-taking that I rely on. Section 4 lays out my empirical strategy and results, and in Section 5 I discuss the implications of my results for HIV prevention policy and health risk information campaigns more generally, and addresses some potential limitations of this paper. Section 6 concludes.

2. Theoretical Framework

This section outlines a model of behavioral responses to HIV risks that relaxes a key assumption made by the previous empirical literature. Most empirical work on responses to risks, HIV or otherwise, relies on an assumption that the stochastic cost of risk-taking is linear in the riskiness of each individual act. My model allows that cost to follow a concave shape that is

[10] See Ahituv et al. (1996) for estimates of HIV risk compensation in the United States and Oster (2012) for Africa.

consistent with the risks of individual sex acts adding up into a sensible total probability of HIV infection. The core result is that the comparative static in question – the derivative of risk-taking with respect to per-act risks – is not always negative, and in general will flip from negative to positive depending on an agent's risk beliefs and risk-taking time profile. In Section 2.1, I lay out the basic form of the model. I focus on the function that aggregates per-act risks into a subjective belief about HIV infection, and in particular on the properties it must satisfy in order to produce sensible probabilities. Using those properties, I then derive the key comparative static, which is the response of sexual risk-taking to the perceived per-act risk of HIV infection (Section 2.2). I show that this response is initially negative (consistent with the previous empirical literature) but becomes positive for sufficiently-high per-act risks.

2.1. Model Basics

In this model, I assume that people weigh the benefits of choosing a level of risky sex, y, against its costs. These costs include both a fixed per-act price (or time cost, or emotional cost) q, and a stochastic component due to the risk of HIV infection. An agent's *perceived* risk per sex act or "riskiness" is x. The expected cost of HIV infection is the agent's subjective belief about the total probability of it occurring, P, times its perceived cost, c. The subjective probability can be written as a continuously differentiable function $P = P(x, n)$, where $n = y + m$ is the total number of sex acts, including both the current choice y and an immutable stock of acts m. This stock includes all previous sex acts since one's most recent HIV test, and also all future risky acts that are unavoidable. The latter captures accidental exposures through things like condom breakage as well as situations where an agent may lack the bargaining power to turn down some future sex acts. It is possible to compute the actual total probability using the binomial distribution, but my results will be robust to agents potentially not understanding how to correctly compute probabilities. The benefit of y sex acts is described by a continuously differentiable benefit function, $B(y)$, with positive and diminishing marginal benefits. This yields the following optimization problem:

$$\max_{y \geq 0} \{U(y; x, m, q, c)\} = \max_{y \geq 0} \{B(y) - qy - P(x, y + m)c\} \qquad (1)$$

By the assumption that *y* is continuous, the maximand $U(n; m, p, c, r)$ is the sum of continuously differentiable functions and therefore continuously differentiable itself.

I do not assume that agents can correctly convert levels of risk-taking and per-act risks into an aggregate probability of HIV infection. Instead, I simply assume that $P(x, y + m)$

corresponds to sensible probabilities: it must lie between 0 and 1, and be equal to zero if either sex is risk-free ($x = 0$) or an agent engages in no risky sex ($y + m = 0$). I also assume that higher riskiness x is in fact interpreted as leading to a higher subjective probability of HIV infection, and more risk-taking $y + m$ also increases the chance of contracting HIV.[11] The subjective probability also approaches 1 as riskiness rises toward 1 and total risk-taking goes to infinity.[12]

For most possible functional forms of $B(\cdot)$ and $P(\cdot, \cdot)$ this optimization problem will have no closed-form solutions for the optimal number of sex acts y^*. However, there must be *some* interior solution as long as the marginal benefit of risky sex outweighs the costs for at least one act, and approaches zero as $y \to \infty$. A necessary condition for interior optima is that $q > 0$, so there is some fixed price or time cost to risky sex.[13] In Section 2.2 I discuss the properties of interior solutions for y^* using the implicit function theorem.

2.2. Comparative Statics

For an interior solution the optimal number of sex acts y^*, must satisfy the following first- and second-order conditions:

$$B'(y^*) - q - P_2(x, y^* + m)c = 0$$
$$B''(y^*) - P_{22}(x, y^* + m)c \leq 0$$

The first-order condition is equivalent to there being a function $G(y^*, x, m, q, c) = B'(y^*) - q - P_2(x, y^* + m) * c = 0$. Therefore the implicit function theorem allows us to compute the comparative static for changes in y^* in response to changes in x:

$$\frac{\partial y^*}{\partial x} = -\frac{\frac{\partial G}{\partial x}}{\frac{\partial G}{\partial y^*}} = \frac{P_{21}(x, y^* + m)c}{B''(y^*) - P_{22}(x, y^* + m)c}$$

The denominator is just the left-hand side of the second-order condition, and is thus negative.[14] Since $c > 0$, $\text{sign}(\frac{\partial y^*}{\partial x}) = -\text{sign}(P_{21}(x, y^* + m))$. As noted in Section 1, it is typical to approximate P by a linear function, $P(x, y + m) \approx x(y + m)$. (This is done explicitly in Oster

[11] Formally, $P_1 \geq 0$, with $P_1(0, y + m) > 0$ if $y + m > 0$ and $P_1(x, 0) = 0$; $P_2 \geq 0$, with $P_2(x, 0) > 0$ if $x > 0$ and $P_2(0, y + m) = 0$.

[12] $P \to 1$ as $y + m \to \infty$ as long as $x > 0$, and $P = 1$ if $x = 1$ and $y + m \neq 0$.

[13] See Online Appendix IA for a proof.

[14] Technically it is only weakly negative since the second-order condition is a weak inequality. The discussion that follows assumes strict negativity, since otherwise $\frac{\partial y^*}{\partial x}$ is undefined. However, all the results in this section will hold as the second-order condition approaches 0 from above.

(2012) and implicitly by Viscusi (1990), for example). In this case $P_{21} = 1$ and $P_{21}(x, y^* + m) > 0$ always, so $\frac{\partial y^*}{\partial x} < 0$. This is the equivalent to the Oster (2012) result that sexual activity should fall as the prevalence of HIV rises. More broadly, it says that behavior is uniformly self-protective: people always choose fewer risky acts as the per-act riskiness of each act rises.

However, the linear approximation specified for the functional form of $P(x, y + m)$ does not, in general, satisfy the requirements for being a sensible probability laid out in Section 2.1. For low values of x and $y + m$ this is not an issue, since P will lie between 0 and 1. In the context of HIV risk beliefs, however, x is often quite high, since perceived risks are typically large overestimates, and m will reflect a potentially long sexual history and an extensive future of possible condom failures and so forth. One way of imposing that probabilities are sensible is to use the true probability function $P = \pi(x, y + m) = 1 - (1 - x)^{y+m}$. O'Donoghue and Rabin (2001) point out that for this function, $\pi_{12} = (1 - x)^{y+m-1}[1 + y + m \ln(1 - x)]$, and hence $\pi_{12} > 0$ if $y + m < \frac{1}{-\ln(1-x)}$ and $\pi_{12} < 0$ if $y + m > \frac{1}{-\ln(1-x)}$. In other words, P_{12} is not constant in sign, but shifts from positive to negative if x rises above a point defined by the total number of risky acts $y + m$. This then implies that the sign of $\frac{\partial y^*}{\partial x}$ will shift from negative to positive when it crosses that tipping point.[15]

This is not specific to relying on the true function $\pi(x, y + m)$ but is true for any function $P(r, y + m)$ that satisfies the basic conditions laid out in Section 2.1. This fact is proven formally in Section IB of the Online Appendix, but can readily be understood from the conceptual illustration in Figure 2. The horizontal axis shows the number of risky acts chosen, while the vertical axis shows the total subjective probability of contracting HIV. The blue line shows the relationship between P and $y + m$ for a low perceived per-act risk x, and the red line shows the relationship for a higher value of x. Consistent with the basic rules of sensible probabilities, and also with the linear approximation used in most empirical research on risk responses, the slope of the red line is initially higher. When sex is riskier, the total probability of contracting HIV initially rises faster for the same number of sex acts. But the total probability is

[15] Sterck (2012) embeds an identical logic in a dynamic framework, assuming that people understand how probabilities actually add up but allowing them to occasionally make mistakes, which drive them into fatalism.

capped at one, so there must be some point above which the slope of the red line is *lower* than that of the blue line.[16] Formally, this can be written as follows:

Proposition 1

$$\exists \tilde{x} = \tilde{x}(y+m) \text{ s.t. } P_{12}(x, y+m) > 0 \text{ if } x < \tilde{x} \text{ and } P_{12}(x, y+m) < 0 \text{ if } x > \tilde{x}$$

Recall that part of the total level of risk taking is tied up in m, which is out of the agent's control. It is useful to think about this as including the agent's sexual history (in a context where HIV testing is unavailable, for example), but it also contains all future risks that the agent cannot avoid. To fix concepts, suppose that everyone thinks that they will experience at least one condom break some time in the future, so $m \geq 1$. For $m = 1$, and using the true function $\pi(x, y+m)$, the tipping point occurs at $x = 0.63$. This is extremely high compared with the actual per-unprotected-act risk of contracting HIV from a randomly-selected partner, but it is not particularly high compared with the subjective beliefs expressed by people in Malawi. At baseline, 28% of my sample believed the risk was at least that high.

If I maintain the assumption that sexually active adults cannot eliminate all possible exposures to HIV (so $m \geq 1$ in general), this eliminates the possibility of a corner solution where $y + m = 0$, and guarantees that the tipping point value $\tilde{x}$ that changes the sign of P_{12} to negative will be somewhere below 1. Proposition 1 then implies that $\frac{\partial y^*}{\partial x}$ will itself have a tipping point:

Proposition 2: Comparative static of *y* with respect to *x*

$$\exists \tilde{x} = \tilde{x}(y+m) \text{ s.t. } \frac{\partial y^*}{\partial x} < 0 \text{ if } x < \tilde{x} \text{ and } \frac{\partial y^*}{\partial x} > 0 \text{ if } x > \tilde{x}$$

Below the threshold value of the per-act HIV infection risk $\tilde{x}$, rational agents will behave self-protectively (responding negatively to risks); above $\tilde{x}$ they will behave fatalistically (responding positively to risks).

This result is somewhat counterintuitive, but it captures a fairly simple logical conclusion: if the risks are sufficiently high and I can't totally avoid exposure, there is no value to limiting how much sex I have. This sort of fatalistic response is a potential issue for a wide range of health decisions. Anti-smoking campaigns, to take one example, often feature "Benefit Timelines" that emphasize the health benefits that accrue to ex-smokers 20 minutes after

[16] The results here technically rely on $P(x, y+m)$ being continuous, but as discussed in Section IC of the Online Appendix it is possible to reach similar conclusions even if people use heuristic methods for aggregating risks into total probabilities that are not continuous.

quitting, 24 hours, 3 months, and so forth.[17] These timelines can be understood as a way to combat the possibility that smokers will think they are doomed to eventual cancer, no matter what they now decide. Similar to the benefit timelines in logic, HIV messaging targeted at HIV-positive people emphasizes the risk of "reinfection" with a different strain of HIV.[18] Actual cases of reinfection are rare enough that the medical importance of this possibility is unclear (Smith et al. 2005), but one goal of this kind of messaging is to avoid a rise in risky sex by selfishly rational people who believe they have nothing to lose. There is also suggestive empirical evidence of fatalism in the African HIV epidemic. Wilson et al. (2012) find that men in Kenya reduce their levels of sexual activity after being circumcised. The authors suggest that this is consistent with a decline in fatalism, but without direct evidence on the men's beliefs about the transmission rate of the virus, they cannot test this theory.

It is possible to extend this result in Proposition 2 to account for altruistic behavior on the part of people who are HIV-positive, who may choose to be careful to protect their sex partners. In this case, there is no stochastic personal cost of risky sex, and $P(x, y+m)$ can instead be interpreted as the total subjective probability of infecting one's partner given a perceived risk x and total risk-taking $y+m$. c is then the extent to which agents care about their partners avoiding HIV. All the same results then go through: for relatively low values of perceived risks and low levels of risk-taking, agents will respond to rises in the per-act risk by reducing how much sex they have, but when the risks are sufficiently high they give up, assuming their partner is either already infected or doomed to infection in the future.

One consequence of Proposition 2 is that the linear relationship between x and y typically estimated in empirical analyses of risk responses may be misspecified, since y is a non-monotonic function of x. Estimated average marginal effects of x on y will in general include both positive and negative ranges of $\frac{\partial y^*}{\partial x}$, which will tend to push the average toward zero. They will also ignore potentially-crucial heterogeneity in the effect of risk beliefs on risky behavior. In my empirical analysis in Section 4, I will explicitly examine risk responses for heterogeneity by initial beliefs.

[17] See e.g. the "Quitting Timeline" from the British National Health Service (2013).
[18] For example HIV Reinfection - Positive Prevention - Reinfection (2007).

3. Data and Experimental Design

This section outlines the data and experiment that I use to test the model laid out in Section 2. I begin by describing the randomized field experiment that I conducted in Southern Malawi to collect data on how individuals' sexual behavior responds to changes in their beliefs about HIV infection risks. I then describe my preferred of measures of sexual risk-taking, which come from retrospective sexual diaries conducted as part of the survey. Finally, I discuss my measures of beliefs about HIV infection risks.

3.1. Experimental Design

This paper uses data from a field randomized controlled trial (RCT) I conducted from August to December 2012. The experiment took place in Traditional Authority (TA) Mwambo, in the Zomba District of Malawi's Southern Region. I sampled roughly 30 sexually active adults aged 18-49 from each of 70 villages. Each participant was interviewed twice: once for a baseline survey, and again for a followup conducted 1-3 months later. At the end of the baseline survey, all participants were provided with basic information about the sexual transmission of HIV and the benefits of condoms.[19] Participants from half of the villages, chosen at random, were also read an information script that presented the actual annual risk of HIV transmission in serodiscordant couples (relationships with one HIV-positive and one HIV-negative partner) that have unprotected sex, based on estimates from Wawer et al. (2005) and also figures from the Malawi National AIDS Commission.

The village sample for the study was constructed from the Malawi National Statistics Office (NSO) GIS files for the 2008 Census. I began by removing all duplicate village entries from the dataset.[20] Because existing evidence indicates that fatalistic responses to HIV risks and risky sexual activity may be concentrated around major trading centers (Kaler 2003), I then constructed sampling strata based on the distance to the closest major trading center.[21] 24 of the

[19] Knowledge of the basics of HIV transmission and prevention is already high in this population. In the 2010 DHS, nearly 100% of individuals said that HIV was sexually transmitted and over four fifths knew that condoms were effective prevention (Malawi National Statistical Office and ORC MACRO, 2010). The latter figure may be an underestimate: in survey questions about the risks of unprotected and protected sex, virtually all respondents stated that condoms provided at least some risk benefit.

[20] The Population and Housing Census uses Enumeration Areas as its basic sampling unit, rather than villages. The boundaries of these enumeration areas commonly cross through villages, leading to duplicate entries.

[21] Trading centers were identified based on the 2008 Malawi Population and Housing Census, which codes peri-urban areas outside the main cities with enumeration area numbers from 800 to 899. I included trading centers both inside the TA as well as in other nearby parts of the Southern Region. Since TA Mwambo adjoins the city of Zomba, I also included the main markets in that city as trading center equivalents. In addition, based on conversations with

sampled villages (34%) were within 2 km of a trading center[22]; another 24 (34%) were within 2 and 5 km from a trading center; and 22 (31%) were more than 5 km away from the closest center. This compares with overall proportions of 10%, 40% and 50% of all villages in TA Mwambo. Within each sampling stratum, I randomly assigned half of the villages to the treatment group and half to the control group.

In each village, a team of enumerators first conducted a comprehensive household census. Using this census, 15 men and 15 women aged 18-49 were then sampled from each village, with only one respondent allowed per household. The sample was thus stratified by both gender and distance to the nearest trading center, so the effective sampling strata are formed by combinations of gender and distance indicators. Some villages had too few households for 30 eligible-age adults to be selected, and hence the maximum feasible number was chosen instead. This yielded a total of 2024 sampled individuals. The survey team then attempted to contact all sampled people for a baseline survey. Although refusals were rare (<1% of respondents refused the baseline survey), 23% of sampled people could not be found at baseline, typically because they were temporarily away from the household; it is common for people in this area of Malawi to travel during the agricultural off-season to look for casual wage labor. A total of 1543 respondents had a successful baseline survey. Because the survey contained sensitive questions about sexual behavior, and the model of fatalism applies mainly to sexually active adults, the survey contained an early screening question that eliminated people who had never had sex from the sample. This included 2.6% of the respondents, leaving 1503 sexually-active adults in the baseline survey.

After a minimum delay of 30 days, the enumerator team attempted to recontact all 1503 sexually-active respondents from the baseline survey, successfully finding 1292. Table 1 presents the characteristics of the final respondent sample. It is balanced across the treatment and control groups, both by the distance categories and by gender. There is also no evidence of differential attrition: an indicator for inclusion in the final sample is not significantly correlated with treatment status, irrespective of whether I control for other baseline covariates (Table 2); these results are robust to running the regressions as logits instead of LPMs.

---

key informants, several more trading centers in the local area were included, that do not have enumeration area codes between 800 and 899 but that are nonetheless major centers for trade and nightlife.

[22] In discussions with key informants in TA Mwambo, 2 km was generally agreed to be the maximal distance people will walk for nightlife. These strata thus roughly proxy for how easily people could access the trading centers in order to drink and search for sex partners.

Baseline summary statistics for the overall sample, as well as a comparison of the treatment and control groups, are presented in Table 3.[23] The sample is 43% male and 82% married, with a mean age just below 30. Respondents are fairly poor on average: household cash expenditures average just under $2 (at purchasing-power parity) per person per day. The sample is well-balanced across the treatment and control groups with the exception of household cash income, which is approximately $64/month higher in the control group. However, this discrepancy can be largely attributed to seasonality in income combined with the differential timing of the baseline surveys (for reasons discussed in Section 3.1.1 below, the control group baseline surveys were done first and the treatment group baseline surveys were done second). A comparison of incomes at the followup survey is valid if we make the plausible assumption that the information treatment had no impact on earnings. Monthly household income at the followup survey is still $23 higher than in the control group, but this difference is not statistically significant. The summary statistics are consistent with the randomization having successfully generated balanced treatment and control groups.

3.1.1. Information Treatment

At the end of the baseline survey, all respondents from the treatment villages were read and shown information about the true risk of HIV infection between serodiscordant partners who have unprotected sex, as measured by the Wawer et al. (2005) study of serodiscordant couples in Rakai, Uganda. The information treatment was administered by the survey enumerators in a one-on-one setting. It involved both an oral component and an interactive visual component. In the oral component, the basic details of the original study were explained, with certain aspects simplified for clarity. Respondents were told that the study occurred in Uganda, and that 100 serodiscordant couples were followed for a single year.[24] They were told that all the couples had regular sex without using condoms, about once every three days on average, and asked how many people they thought would contract HIV. They were then informed that in fact only ten of

[23] In this table, and in all the other balance tests in this paper, the p-values are adjusted to account for the clustered design of the study, following Donner and Klar (2000).

[24] The actual figure for the Wawer et al. study is 235 couples, 188 of which never used condoms when they had sex (results are not broken out by condom use, but condom use was very inconsistent and had no impact on the estimated transmission rate). The time period was actually 10 months, with some couples being observed for multiple time windows. This was reduced to 100 couples over the course of 1 year for the sake of clarity and simplicity.

the initially HIV-negative people became HIV-positive.[25] Respondents were asked if they believed the results of the study, and enumerators were trained in how to respond to a number of common questions, such as whether the testing equipment was faulty. The script listed the reasons that HIV transmission sometimes does not happen even when serodiscordant couples have unprotected sex, for example the fact that HIV sometimes cannot penetrate the genitalia. It is then emphasized that HIV transmission is something that happens by chance, comparing it to a popular games of chance used by local cell phone companies as marketing tools.

The interactive visual component complemented the oral component and occurred at the same time. It involved showing respondents a diagram with 100 pairs of stick figures representing serodiscordant couples, with a black stick figure indicating an HIV-negative partner and white stick figure indicating an HIV-positive partner. The respondent was asked to guess as to the number of people who would contract HIV after a year of regular unprotected sex with an infected partner, and this guess was indicated by circling an appropriate number of these stick figure couples. When the true rate was presented, the enumerator showed a second diagram in which ten of the initially HIV-negative individuals had turned from black to white. Enumerators then counted and circled these transmissions.

To minimize the risk of contaminating the control villages, all the baseline treatment surveys were done after the baseline control surveys were completed. This approach was based on Godlonton et al. (2012). The survey enumerators were only taught to administer the information intervention after all the control surveys were completed.

3.1.2. Potential ethical concerns

The key potential ethical concern about the design of this study was that people may respond self-protectively to HIV infection risks on average. In this case the information treatment would increase the average amount of risky sex people have, leaving people in the treatment group worse off. This concern is mitigated by four factors. First, to the extent that we believe responsible adults can be trusted to make their own choices with the information they have, it is appropriate to provide people with better information rather than worse. The *de facto* policy in Malawi is to overstate HIV transmission risks. This strategy is potentially at odds with

[25] This is the annual transmission rate cited by the Malawi National AIDS Commission. The exact annual rate implied by the Wawer results is 12%. The Hollingsworth et al. (2008) reanalysis of the Wawer data finds an annual transmission rate of 10.6% from asymptomatic partners (HIV-positive sex partners who have not just recently contracted the virus and do not yet have AIDS), which are the majority of cases, but does not provide an overall average.

the first ethical principle emphasized in the Belmont Report, which is that individuals should be respected as autonomous persons:

> To respect autonomy is to give weight to autonomous persons' considered opinions and choices while refraining from obstructing their actions unless they are clearly detrimental to others. To show lack of respect for an autonomous agent is to repudiate that person's considered judgments, to deny an individual the freedom to act on those considered judgments, or to withhold information necessary to make a considered judgment, when there are no compelling reasons to do so. (Office of the Secretary, 1979)

Hence the policy of denying people information about the true risks they face is potentially unethical, given that there is very little empirical evidence that would provide compelling reasons to withhold that information.

Second, the information provided to the treatment group is medically-accurate, publicly available information. It is also the same information provided by the Malawi National AIDS Commission (NAC) in their policy documents, which state that the annual risk of HIV transmission is 10% (Malawi National AIDS Commission 2003, p. 11). NAC's official policy is also that HIV information and education programs should provide accurate information about safer sex (Malawi National AIDS Commission 2003, p. 6).[26] Hence the additional information provided to the treatment group is completely consistent with Malawi government policy, and can be seen as a test of what would happen if HIV information and education campaigns actually provided HIV transmission risk information that is consistent with what NAC provides on its website.

A third mitigating factor is that previous estimates of responses to HIV risks in Africa are very small in magnitude (e.g. Oster 2012), and the *ex ante* expected impact of the information treatment was small, limiting any potential harm. The reason that the experiment was still interesting was that the responses were not expected to be uniform. There is reason to believe that many people in Malawi may react fatalistically to HIV risks. As mentioned above, cross-sectional data from elsewhere in Zomba District shows suggestive evidence that the response of

[26] The exact quote is:
"Government, through the NAC, undertakes to do the following:
• ensure that all people have equal access to culturally-sound and age-appropriate formal and nonformal HIV/AIDS information and education programmes, which shall include free and accurate information regarding mother-to-child transmission, breastfeeding, treatment, nutrition, change of lifestyle, safer sex and the importance of respect for and nondiscrimination against PLWAs"

sexual behavior to HIV infection is positive for people with high risk beliefs (Kerwin 2012 WP). Kaler (2003) documents men from rural Southern Malawi employing fatalistic reason - saying that it is sometimes not worthwhile to use condoms, because the risk of contracting the virus is so high:

> And then I asked my in-law, "What do people do after noticing that his/her partner seems to have AIDS?" He said, "Some couples come to an end and for others the marriage continues." And I asked, "Do they use condoms then?" He said "I don't think they use [them] because it will just be a waste of time since both of them have contracted the disease." (Simon, journal May 3 2002)

For people who respond fatalistically, learning that their assessment of the risk is an overestimate will actually reduce sexual risk-taking, rather than increasing it. This experiment was designed to capture heterogeneity in responses around a mean response that is small in magnitude.

Finally, this concern is mitigated because excessively high risk beliefs may have negative long-term effects independent on direct effects on sexual behavior. As people realize that it is possible to for sexually active married couples to remain serodiscordant for a long time, they may lose trust in the medical and science community or the education system, and may also promulgate false rumors about HIV transmission and immunity. Since most people believe that the transmission rate of HIV is 100%, they may instead falsely assume that continued serodiscordance means that a specific person or group is immune to the virus. There is already evidence that the latter is going on: 42% of my respondents said that they believed people with type-O blood were immune to HIV, an idea which has no basis in scientific fact.

A separate potential concern is that the information presented is about the approximate overall average risk, but transmission risks actually vary by demographic groups. For example, the transmission rate is 3 to 5 times higher for women than for men, and about 60% lower for circumcised men than for uncircumcised men. However, this concern is mitigated by the fact that baseline beliefs are very high (93% per year on average for the control group). Hence virtually all respondents in the treatment group have more-accurate beliefs after the information treatment than they did beforehand.

To ensure that respondents' well-being was protected, ethics oversight for this study was provided by both an in-country IRB (The University of Malawi's College of Medicine Research and Ethics Committee, or COMREC) and one at my home institution (The University of

Michigan's IRB-Health Sciences and Behavioral Sciences, or IRB-HSBS). The final study protocol, including the information treatment, was reviewed and approved by both IRBs. The approved protocol also included a management plan under which preliminary results were provided to the two IRBs in order to manage any possible rise in HIV transmissions as a result of the information treatment. As I discuss in Section 5.2, the actual results of this study suggest essentially no increase in HIV transmissions among the treatment group relative to the status quo.

3.2. Measures of sexual behavior

My primary outcome measure is self-reported sexual behavior as recorded using a retrospective sexual "diary" that walks respondents back through the previous seven days to collect data about a range of activities, focusing specifically on sex. It then asks detailed questions about each reported sex act. In this subsection I begin by discussing the relative merits of self-reports and the main alternative measure used in the literature, STI incidence. I then describe the advantages of the sex diary approach relative to the alternative of single-question recall methods.

3.2.1. Self-reported risky sex versus STI incidence

By relying on self-reports, my approach contrasts with a strand of the HIV prevention literature that relies on STI incidence as an objective measure of risky sexual behavior, since any person with an incident (new) STI has necessarily had unprotected sex. One reason to prefer this approach is that an intervention may increase the social desirability of a behavioral outcome without actually affecting the underlying behavior (Özler 2013). Put simply, if one is studying an intervention that teaches people that unprotected sex is bad, it is likely that people who receive the intervention will report less unprotected sex, even if there is no change in their actual behavior. Measuring the incidence of treatable STIs is advanced as a solution to this problem. If all subjects are tested for treatable STIs at baseline, and treated for those they have, any STI infections measured at endline are proof positive of unprotected sex.

This approach is subject to three important limitations, however. The first is that it can only detect risky sex with an STI-infected partner, and only in cases where a transmission actually occurs. As a result it will tend to understate all risk-taking. This downward bias in the outcome measure will be balanced across study arms, so that all statistical tests will be asymptotically valid, but measured effect sizes will be much smaller than the true effect on risky

sexual behavior. In turn, this will tend to bias finite-sample significance tests toward the null hypothesis due to lower statistical power. Indeed, this concern is one of the reasons that HIV incidence itself is rarely used as an outcome measure: since its baseline incidence is so low, detecting changes requires massive sample sizes.

The second issue with using STIs is that it relies only on new STI infections. The gold standard for the STI outcomes approach is to test all respondents at baseline, and treat everyone who tests positive for curable STIs (syphilis, gonorrhea, and Chlamydia). Incident infections of these STIs then form the outcome measure (Fishbein and Pequegnat 1999). This is the method employed by the de Walque et al. (2012) study of using conditional cash transfers to incentivize safer sex, for example. By relying on treatable STIs and clearing the sample population of them at baseline the researcher can be sure to capture risky sex by all subjects. This is the second motivation for not generally relying on the incidence of HIV itself: we are interested in the behavior of people who already have HIV, and as HIV is not curable we could not detect changes in the behavior of people who already have the virus at baseline if we used HIV incidence as an outcome measure. However, many studies either do not rely on curable STIs (Baird et al. 2012) or do not treat subjects' curable STIs at baseline (Gong 2012). When this is done, risky sex can only be measured among respondents who are initially STI-free. This not only biases estimates toward zero, but omits a crucial group for the HIV epidemic. The initially infected subpopulation represents a key epidemiological group for both biomedical and behavioral reasons. Biomedically, they are at increased risk of contracting and spreading HIV because other STIs increase the HIV transmission rate (Galvin and Cohen 2004). Behaviorally, we have a clear-cut signal of ex ante risky behavior, by the same logic that motivates the STI incidence outcome measure. As sexual behavior is highly autocorrelated, they are also a likely high-activity group for future sex, and hence measuring effects on their behavior is vitally important. This concern is amplified in the case of my study, as theory predicts heterogeneous responses across the population by baseline beliefs. That pattern is likely to correspond to variations in baseline sexual risk-taking, and indeed I do find that people with higher baseline risk beliefs have more lifetime sex partners.[27]

[27] This may help explain the difference between the results Thornton (2008) and Gong (2012) find for the effect of HIV testing on sexual behavior. In contrast to Thornton, Gong finds that people who are surprised by learning that they are HIV-negative have less risky sex, and those who are surprised by learning that they are HIV-positive have more, which is consistent with a selfishly rational net response to the new information.

The third drawback is that the STIs used as outcome measures must be selected carefully: they must measure only the risky activity in question, and must be applicable to the context in question. HSV-2, for example, can be transmitted even if a condom is used during vaginal sex (Holmes et al. 2004), and can be transmitted via oral sex (Lafferty et al. 1987). The latter is a concern if the goal is to measure HIV risk activities, as the best evidence indicates that unprotected oral sex is not a risk factor for HIV transmission.[29] Hence HSV incidence will tend to measure all sexual activity instead of unprotected vaginally sex, which is potentially not a desirable measure from the perspective of HIV prevention. Local areas also vary widely in terms of the prevalence of various STIs, so that the downward bias on the effect size may be more severe in some cases than in others.

In addition to the limitations of STI measures, the drawbacks of self-reports are far more limited for this study than for much research on HIV prevention. Social desirability bias in individual responses will only confound estimates of the impact of a treatment if it is correlated with treatment status. This is typical for sexual behavioral change interventions, in which respondents in the treatment group are informed as to the "correct" answer. However, the information treatment I employed does not include any guidance as to how people should behave. Thus any social desirability bias will tend to attenuate all self-reports of undesirable behaviors equally across the treatment and control groups. This will tend to work against my finding treatment-control differences, making any estimated differences lower bounds on the actual effect size. Moreover, much of my analysis focuses on heterogeneity in responses to the treatment by baseline risk beliefs, with an emphasis on some groups responding negatively to the treatment and some positively. This kind of heterogeneity would not be expected as a result of social desirability bias.

3.2.2. Measures of self-reported sex

The baseline and endline surveys contained two broad categories of self-reported sexual behavior. The first were single-question recall measures, for example: "In the past 30 days, how

[29] No study has ever conclusively documented an HIV transmission due to oral sex alone without either another exposure, or another risk factor such as open sores in the mouth. The highest-quality studies on oral sex and HIV transmission find negligible transmission rates. See Kerwin et al. 2011a for a review of the epidemiological evidence on this topic.

many total times did you have sex, including serious and non-serious partners?" The second comprised a detailed retrospective sexual diary, which walks respondents through the previous seven days beginning with yesterday. On each day, respondents were asked what time they woke up, how much alcohol they had, whether they were menstruating (or for men, whether their sex partner was menstruating), the value of gifts they gave to or received from their partner,[30] how many times they had sex, and the time they went to sleep. Then, for each reported sex act, they were asked detailed questions such as the time of day, the length of the act, condom use, and whether the sex act was with their primary sex partner or a different partner.

This diary-based approach to measuring sexual behavior was initially developed and refined in previous research on sexual behavior in Southern Malawi (Kerwin et al. 2011b). It builds on research that shows that calendar-based methods reduce recall bias compared with single-question recall methods (Belli et al. 2001). Luke et al. (2009) have found that relationship history calendars improve the quality of responses to questions on sexual behavior, showing that apparent biases due to social desirability effects are smaller. The sex diary approach adapts these insights to a much shorter time frame to assist respondents in the recall of all sex acts over the past 7 days. The improved accuracy of the sex diary over other methods is reflected in the data captured by the surveys. Column 1 of Table 4 shows that the two variables record fairly similar levels of sexual activity. However, for the endline survey, the two variables have sharply different means, at 1.72 sex acts per 7 days on the sex diaries and 1.31 in the single-question 30-day recall variable. This difference in means may reflect reduced recall bias from the sex diary measure.

The distributions of the two variables are also very different. Panel A of Figure 3 shows the distribution of sex acts in the past 30 days from the single-question recall variable. It demonstrates substantial "heaping" at multiples of 5, with large spikes at 5, 10, 15, etc. In contrast, the distribution of sex acts in the past 7 days for the sex diary question (Panel B) has no appreciable heaping whatsoever.[31] Unlike classical measurement error, this kind of heaping in the dependent variable may bias the point estimates from a linear regression. In a set of simple simulations, I find that sufficiently high levels of heaping bias the estimated coefficients toward

[30] The culture of gift-giving in sexual relationships in Malawi is strongly gender-driven: with very few exceptions men give gifts to women and not the other way around.

[31] To improve readability, both histograms omit zero, which would simply show up as a large spike on the left and compress the y-axis.

zero (results available upon request). The reason for this is clear from considering the extreme case, where the heaping is so extreme that all values are collapsed to a single point. While one could not run an OLS regression in this case, it is clear that the effect of any variable on this (mismeasured) outcome is zero. Given the heaping issue, and the fact that the two variables have significantly different means in the endline data, I will rely primarily on the sex diary variables for my analysis.

Table 4 presents summary statistics for all the available measures of sexual activiy in the data. Columns 3 and 4 show the means of my measures of sexual activity for the control and treatment groups respectively, while Column 5 shows the difference between the two. These are generally balanced across the two study arms, with the only statistically significant differences being a lower number of lifetime sex partners ($p<0.05$) and a smaller share of unprotected sex acts in the past 7 days ($p<0.1$) in the control group. All the differences are fairly small in magnitude, but none of the variables has exactly equal means across the treatment and control groups at baseline. This is part of the motivation for controlling for respondents' baseline values of self-reported sex, as described in Section 4.2.

An additional measure of the demand for safer sex comes from the sale of subsidized condoms to respondents that occurred immediately after the endline survey. All participants were given six coins worth five Malawi Kwacha each (30 Kwacha total, or about ten cents). They were then offered the chance to purchase 3-packs of condoms for five Kwacha apiece, or individual condoms for two kwacha. The condoms sold were Chishango, the most popular local brand. While the price represents a sizeable subsidy relative to the sale of Chishango condoms at local stores, the vast majority of respondents who acquired condoms got them for free. When asked about the nearest place to acquire condoms, respondents commonly named heath centers and health extension workers, both of which offer condoms free of charge. This measure was only collected at the endline survey.

3.3. Measures of risk beliefs

The central prediction of the model I outline in Section 2 is that individuals' responses to risk information will depend on their initial perceptions of those risks. A key input for my analysis, therefore, is a quantitative measure of risk perceptions. Due to data limitations, one common strategy for this is to utilize some measure of the true risk. This is common in the existing literature on responses to HIV infection risks in the United States (Ahituv et al. 1996)

and in Africa (Oster 2012, Juhn et al. 2009). However, an emerging literature has shown that it is feasible to collect meaningful data on subjective beliefs about probabilities using surveys, not just in the United States (Lillard and Willis 2001) but also in the developing world (Attanasio 2009). Specifically in Malawi, Delavande and Kohler (2009), henceforth DK, have developed a method of eliciting subjective expectations that they have shown performs very well. Their approach, implemented starting with the 2006 wave of the Malawi Longitudinal Study of Families and Health (MLSFH) relies on training respondents to report subjective beliefs by choosing how many of ten beans to place on a plate indicating a belief that an event will happen (with 0 beans meaning a 0% chance and 10 beans indicating a 100% chance), using common events in Malawi as reference points.

Rather than following DK, I rely on measures of subjective risk beliefs collected using concrete questions about proportions out of a fixed number of people. These are questions of the form "If 100 men, who do not have HIV, each sleep with a woman who is HIV positive tonight and do not use a condom, how many of them do you think will have HIV after the night?" I then divide the reported number by the denominator used to construct a subjective probability. Question E1a in Figure 4 is an example of one of these questions. All the questions were gender-specific: for instance, when men were asked about HIV transmission they were asked about 100 men having sex with an HIV-positive woman, and likewise women were asked about 100 women having sex with an HIV-positive man. Table 5 presents summary statistics for the six HIV risk belief variables that were collected: the unprotected transmission rate (both per-act and annual), the condom-protected transmission rate (both per-act and with a condom), and two questions about the prevalence of the virus: the share of all members of the opposite sex that respondents thought were HIV-positive, and the share of members of the opposite sex that they find attractive.

I use these concrete expectations questions for two reasons. First, the DK approach adds considerably to the logistical complexity of surveys, as well as the time needed to conduct them. The former issue was a substantive concern with regard to this study because the surveys for the treatment group already required teaching respondents about the probability of HIV transmission; adding an instruction session about the probability questions risked overloading respondents.

Second, this concrete style of expectation question has been validated through extensive

use in previous research across a variety of contexts in Malawi, including in urban areas (Chinkhumba et al. 2012) as well as in areas of rural southern Malawi near my study site (Godlonton et al. 2012, Kerwin et al. 2011b). This previous use has found them to be effectively scale-invariant: switching the number from 100 to 1000 or 10,000 they yields nearly the same subjective probabilities on average, and respondents give the exact same answer about 70% of the time (Author's calculations based on Chinkhumba et al. 2012). The questions also perform well in terms of respecting nested probabilities: if the chance of event B occurring includes all possible instances of event A, then respondents should ideally report a weakly higher probability for B than for A. DK emphasize this as one of the major strengths of their approach. My data do not afford many direct comparisons with DK's on HIV transmission and HIV prevalence, because the DK survey instrument did not ask many HIV-related questions that are necessarily nested within one another. One comparison, however, is the per-unprotected-sex-act risk of contracting HIV from an infected partner, compared with the annual risk. In my data, the latter probability was weakly higher 92.2% of the time, whereas this was the case 91.9% of the time in the DK data (Author's calculations based on MLSFH 2006).[32] In addition to performing comparably to the DK measure in terms of nesting probabilities, the concrete probability method also produces similar results in terms of the mean expectation of the risk of HIV transmission: this is 82.8% per act for the control group at baseline using concrete probabilities[33], and 85.9% per act using DK.

3.3.1. Focal probabilities

One potential concern with eliciting subjective expectations is the tendency for probabilities to heap at the "focal" probabilities of 0%, 50%, and 100%. The typical interpretation, cited by DK, is that this heaping reflects a misunderstanding of the question, or simple uncertainty, rather than a true belief. For my specific application, heaping at the endpoints is not a major concern: many people do hold beliefs that, say, the per-act transmission rate of HIV is 100% for unprotected sex, and that it is 0% if you use a condom. I therefore follow the strategy used on the MLSFH by DK and treat these as "real" answers, with no followup

[32] The annual question for DK actually asks about someone who is married to an HIV-positive person, and does not explicitly specify unprotected sex. However, social norms in Malawi strongly proscribe the use of condoms within marriages (Tavory and Swidler 2009) and married couples use condoms just 11.2% of the time in my sample. Repeating this analysis just for people in the MLSFH sample who say there is no chance they would use condoms with their own spouse yields a similar nesting rate of 94.1%.

[33] The measured probability differs for the treatment group; see Section 4.3.2 for details.

prompting. This issue is a concern, however, for 50%. People commonly use 50% (or in my case, report half of the total denominator), when they are simply unsure about the answer. To address this issue, respondents who reported beliefs of 50% were prompted with a followup question about whether they really believed the chance was 50%, or if they were just not sure, which is an approach taken on the Health and Retirement Study's subjective expectations questions (Hudomiet et al. 2010). Building on that approach, respondents who said they were just not sure were then prompted for their best guess. Question E1b in Figure 4 illustrates these followup questions. The measure of risk beliefs I rely on uses the results of the main question, but replaces 50% with the expectation measured in the followup question for people who say they are just not sure, and change their mind when prompted for their best guess.

3.3.2. Enumerator-knowledge contamination of measured beliefs

As described in Section 3.1 above, the enumerators were only trained to provide the information intervention after the baseline interviews for the control group were finished. This was done to minimize any chance of the information intervention contaminating the control group. However, it also meant that this was the first time the enumerators were taught the true risk of HIV transmission. As a result, enumerators brought different beliefs with them into the baseline treatment and control surveys. This had a relatively small but statistically significant effect on the measured beliefs of respondents at baseline. Panel A of Table 5 presents a comparison of HIV transmission and prevalence beliefs measured at baseline for the control group and the treatment group. Recorded responses for the treatment group are systematically lower than those for the control group, for both transmission beliefs and prevalence beliefs, and most of these differences are statistically significant.

Figure 5 shows the daily average recorded risk belief, separately for treatment and control surveys, and including both baseline and followup surveys. The lines show linear time trends fit to the data. One thing that is immediately clear is that the measured difference at baseline is much smaller than the impact of the information intervention. This can be confirmed numerically by looking at Table 6, which presents the effects of the information intervention people's beliefs about the transmission rate of the virus.

There are two potential explanations for this pattern. One is that different knowledge may have lead enumerators to prime subjects differently, possibly even subconsciously. Enumerators

were trained to follow up with probing questions when respondents answered a question with by just saying that they did not know. The phrasing of these probing questions could have been affected by the knowledge enumerators brought to the surveys. A second is enumerator experience with the questions. While the sex diary questions that form my outcome measure use very simple statements that enumerators were already familiar with using, the phrasing used on the subjective expectations questions was fairly complex. This may have lead to some temporal pattern in reported risk beliefs as the phrasing and probing questions used were refined over time.

There is evidence for both explanations in Figure 5. A downward trend in measured risk beliefs is evident prior to the enumerators being taught the information about HIV transmission, and there is a large drop in beliefs after the yellow line that marks the training session. Further confirmation of the importance of enumerator knowledge for measured risk beliefs can be seen based on the light blue dots that appear after the yellow line. These are average beliefs for "cleanup" surveys – a handful of control-group interviews that were done after the treatment surveys had begun, because respondents were not home when surveys were attempted prior to the information treatment training. Excluding the large negative outlier (which is the average for a day when just a single control-group survey was done) these generally match the measured beliefs for the baseline treatment group surveys.

Another way of understanding the importance of enumerator knowledge is to compare the beliefs recorded at baseline for the treatment group to the endline beliefs for the control group. These are both surveys during which the enumerators' knowledge is identical. Panel B of Table 5 shows this comparison. The only statistically-significant differences are in annual unprotected transmission risks and the prevalence of HIV among attractive people of the opposite sex.

To correct for this apparent contamination, I adjust reported beliefs based on time trends with a trend break. This involves running the following regression:

$$x_{i0} = \rho_0 + \rho_1 Date_i + \rho_2 After_i + \rho_3 After_i * Date_i + v_i$$

$Date_i$ is the date of the baseline survey for respondent i and $After_i$ is an indicator for whether the baseline survey was done after the information treatment training session. I then construct $x_{i0}^{Adj} = x_{i0} - \rho_1 Date_i - \rho_2 After_i - \rho_3 After_i * Date_i$. Panel C of Table 5 presents the trend-adjusted risk beliefs for the control and treatment groups. They are unsurprisingly similar across

groups. As robustness checks, I also replicate my analysis using the raw (unadjusted) risk belief measures, as well as two other kinds of trend adjustment: using a single trend across the whole baseline period, and using just a level shift in reported beliefs. My results are not sensitive to any of these variations, but my preferred specifications use the adjustments described above. These have a simple interpretation: they are my best estimate of how a respondent's initial beliefs compare with the rest of the sample, given the known time trend and trend break evident in the data due to enumerator knowledge contamination.

3.3.3. Composite belief measures

I also construct composite measures of the perceived risk of HIV infection from unprotected sex with any partner, rather than one that is specifically HIV positive. This involves multiplying the perceived prevalence of HIV by the perceived risk of contracting the virus conditional on an HIV-positive partner. Consistent with the specifications laid out prior to data collection, in an earlier working paper version of this paper (Kerwin 2012 WP), I focus on the product of the perceived per-act risk of HIV transmission from unprotected sex with an infected part and the perceived prevalence of HIV among attractive people of the opposite gender (in Panel C of Table 5, the first row times the sixth row). I do this for three reasons. First, even though the information intervention was about annual HIV transmission from an infected partner, it affected people's beliefs about all HIV prevalence and transmission risks (Table 6). As I discuss in Section 4.1, we would expect this Bayesian-style updating of related beliefs if people actually understand and learn the information, rather than just memorizing it. Relying on variation in the other beliefs allows us to avoid one of the shortcomings of using perceived annual HIV risks, which is that they are extremely concentrated in the right tail. At baseline, over seven in ten respondents believe that the annual risk of HIV transmission from unprotected sex is 100%. There is considerably more variation in perceived per-act risks and perceived prevalence.

Second, I use perceived HIV prevalence among attractive people of the opposite sex to mitigate concerns about people holding unrealistically optimistic beliefs: they may believe that the general population of women has a higher prevalence of HIV than the people they are interested in having sex with. Weinstein and Klein (1996) document that people have unrealistically optimistic beliefs about their own risks of experiencing a wide range of negative events. There is a parallel concern of unrealistic pessimism: people's stated perceptions of both HIV prevalence and transmission risks are much higher than the truth, and it may be the case that

they feel more at risk personally than they believe to be the case for the broader population. While I cannot in general eliminate the potential issue of unrealistic optimism, focusing on attractive members of the opposite sex (rather than all local people of the opposite sex) is likely to be a superior measure of the level of risk people feel they actually face. Third, using the product of per-act transmission risks and HIV prevalence gives a natural measure of how risky a respondent believes unprotected sex to be. While this will not apply to any specific sex partner or prospective sex partner, it does capture the general level of HIV infection risk that respondents believe they face when having unprotected sex. If people think that the transmission rate of HIV is high, but that almost none of their prospective sex partners is infected, then the effective perceived risk they face is in fact quite low.

In Figures 6 and 7 I present baseline histograms of the risk measure I will focus on in this paper, constructed two different ways. Figure 6 uses unadjusted values of the per-act risk and prevalence belief variables (from Panel A of Table 5), while Figure 7 uses values that have been adjusted for a linear time trend with a trend break (taken from Panel C of Table 5). The distributions are broadly similar with and without the adjustment. One thing to note is that I do not bound the trend-adjusted beliefs to lie within [0, 1], so the left and right tails of the histogram in Figure 7 go slightly above zero and one respectively.[34]

4. Empirical Strategy and Results

This section details the empirical strategy I employ in analyzing my data. I begin by showing that the information treatment had large effects on people's risk beliefs. I then disucss the optimal estimator to measure the impact of the treatment, which involves controlling for baseline values of the outcome variable. I also lay out the two different combined sexual risk activity indices that I use in my analysis – one that uses all available outcome measures, and another that focuses on the more accurately-measured outcomes from the sex diary. Using these, I show that the average effect of the information treatment is to slightly (but statistically significantly) increase the amount of risky sex people have. This is consistent with a small negative risk elasticity of risky sex, which I estimate directly using two-stage least squares.

I employ the conventional heterogeneous treatment effects analysis strategy of interacting the treatment group indicator with baseline covariates to show that the overall average masks

[34] My main results are robust to recoding negative beliefs to be zero and beliefs above one to be one.

substantial heterogeneity by baseline beliefs, but not by other baseline variables. I then extend this same analysis to marginal effects as well: I use indirect least squares to develop an estimator of the local average treatment effect (LATE) that allows for heterogeneity by baseline covariates. Using this heterogeneous LATE estimator, I show that the marginal effect of risk beliefs on risky sexual behavior is negative for individuals with low risk beliefs, and becomes positive for individuals with at the high end of the risk belief distribution. I show that the sample-average LATE is an average of these heterogeneous LATEs, weighted by the degree of compliance for each value of the baseline covariate.

4.1. Impact of the information treatment on risk beliefs

The information treatment had large effects on respondents' risk beliefs. Table 6 shows the endline treatment-control differences for all the measures of people's beliefs about HIV transmission and prevalence. The treatment group believes the annual risk from unprotected sex is 38 percentage points lower than the control group does. Their belief about the per-act risk has decreased even further, by 43 percentage points. This discrepancy can be understood as a consequence of the ceiling of 100% on transmission rates; 50% of treatment group respondents who think the annual transmission rate is 100% believe the per-act transmission rate is less than that. Note that the respondents do not update their beliefs perfectly: the actual annual transmission rate is about 10%; just 2% of the sample reports beliefs low enough to be consistent with the true values (given that the questions have a minimum granularity of 1%).

Respondents also update their beliefs about HIV risk variables other than the transmission rate from unprotected sex. For example, beliefs about condom-protected sex and about prevalence are both reduced. This suggests that more than simply memorizing the numbers they were told, respondents learned the information and updated their beliefs accordingly. If people understand that the current prevalence of HIV depends on infected people transmitting the virus to others, then a reduction in the transmission rate implies the prevalence of the virus should be lower as well. The information treatment contained no direct information about the prevalence of the virus or on condom-protected sex, so the effects on these variable can be ascribed purely to this learning process. Another factor suggesting that respondents learned the new information and updated their beliefs accordingly, rather than simple memorization, is that risk beliefs were lower even for individuals who did not report an annual transmission rate of

10% or less. At endline, the mean among annual transmission rate belief for this group was 70.5% for the treatment group, versus 90.7% for the control group.

4.2. Choice of Estimator

My estimates of the effect of the treatment on sexual behavior rely on controlling for baseline values of the outcome variable, an approach developed by Frison and Pocock (1991) and McKenzie (2012) and referred to as the "ANCOVA" estimator for historical reasons. They show that when baseline data on outcomes exists, this method has lower variance than relying the endline values of the outcome alone (the "POST" estimator) or using differences-in-differences (the "DIFF" estimator, referred to as a value-added approach in the education literature). By extending Frison and Pocock's analysis, one can also show that the ANCOVA estimator is less biased than the alternatives when there are baseline differences in outcomes/ (See Section II of the Online Appendix for a mathematical derivation). Note that this bias-minimizing property will be relevant whenever there is any non-zero baseline difference in outcomes, even if that difference is not statistically significant, and so this is the optimal estimator to use in almost any finite sample.

The specifications used in this paper will also control for the stratification cells (combinations of distance categories and gender) used to draw the original sample, which can be shown to improve statistical efficiency (Bruhn and McKenzie 2009). Hence my regressions will have the following form:

$$y_{i1} = \alpha + \beta_{ANCOVA} T_{i1} + \gamma y_{i0} + \eta Z_{i0} + e_i$$

where Z_{i0} is a vector of categorical dummy variables for the sampling strata.

4.3. Combined outcome indices

It is common in the literature present results using a constructed combined outcome index, both to reduce concerns about multiple comparisons and to improve the precision of estimates (e.g Godlonton et al. 2012, Kling et al. 2007). However, the value of such an index is

unclear in situations where some outcomes are measured with greater error or where baseline data is not available for particular outcomes (for example, condom sales were only done at endline). I therefore present two versions of the sexual risk index. One uses all outcomes that can be constructed from the retrospective sexual diary, which I argue in Section 3.2 above provides more accurately-measured outcomes than the single-question recall variables that comprise my other available- outcome variables. An alternative index includes both the sex diary outcomes as well as all other outcomes that can be constructed from the survey.[36]

Each index is constructed separately for the baseline and followup waves by normalizing all component variables (subtracting the sample mean and the dividing by the sample standard deviation). The normalization is reversed in sign for condom use, condom acquisition, and condoms purchased, for which positive numbers imply less risk-taking. These normalized values are then averaged for each respondent. The indices are constructed ignoring any missing data, so if a respondent did not answer one of the questions (or if it was not applicable, for instance share of unprotected sex acts in the past week for someone who did not have any sex) the index is constructed using just the variables that do have valid data.

4.4. Intent-to-treat effects

The results of the intent-to-treat specifications are shown in Table 7. The estimated impact is small in magnitude: it is possible to rule out magnitudes larger than 0.25 standard deviations. For the best-measured outcomes, from the retrospective sex diary, the effects are statistically significant at the 10% level (although still small). The number of sex acts in the past week rises by 0.2, which is 12% of the pooled baseline mean, or about 0.09 standard deviations. Focusing specifically on the margin of abstinence (whether people have any sex at all), this shifts by 5 percentage points, which is roughly 0.1 standard deviations. The overall sex diary risk index

[36] The diary sexual risk index includes six variables recorded on the sex diary: any sex in the past week, total sex acts in the past week, share of sex acts unprotected in the past week, sex with more than one partner in the past week, total sex acts with non-primary partners in the past week, and the share of sex acts with non-primary partners that are unprotected. The overall sexual risk index includes all the outcomes from the diary sexual risk index, as well as any sex in the past month, total sex acts in the past month, share of sex acts in the past month that are unprotected, total sex partners in the past 30 days, condoms acquired in the past 30 days, any condoms bought at the time of the survey, and total condoms bought at the time of the survey.

also rises by 0.07 SDs.

The treatment has no effect on condom use, nor on condom purchases. There is a significant effect on the number of condoms acquired (either bought or gotten for free) in the past 30 days, but this result on condom acquisitions is fairly sensitive to outliers.[37] This is consistent with the extremely high rates of unprotected sex: at baseline just 1 in 10 sex acts involved a condom, leaving limited room for increases in risk-taking at this margin.

While the results in Table 7 represent my preferred specifications for looking at the intent-to-treat effect of the information treatment, it is possible to defend a number of alternative specifications for each outcome. For example, some of my outcomes are technically count variables, and so a negative binomial regression may be preferable to ordinary least squares. A sensitivity analysis shows that the estimates in Table 7 are robust to a variety of different specifications; the t-statistics for tests of differences between the preferred marginal effect estimates and the alternatives are uniformly below one, and most of the estimated mean marginal effects are very similar.[38] I therefore rely on the specifications described in Section 4.1 for the remainder of my analysis.

4.5. The marginal effect of risk beliefs on risky sex

More generalizable than the specific effect of this information treatment on sexual behavior is the marginal effect of HIV risk beliefs on sexual risk-taking. As above, my measure of risk beliefs in this section is the perceived chance of contracting HIV from one unprotected sex act with a randomly-chosen attractive person from the local area.[39] Specifically, the regression I estimate is

$$y_{i1} = \alpha + \delta_{OLS} x_{i1} + \gamma y_{i0} + \eta Z_{i0} + e_i$$

[37] The analysis in Table 7 top-codes condoms acquired at 144 (four boxes of Chishango condoms), which affects solely respondents in the treatment group at endline, and hence reduces the estimated effect relative to the results without top-coding. Topcoding condom acquisitions at 30 reduces the estimated effect to 1.13, with a standard error of 0.46, which is still significant at the 0.05 level. However, excluding all observations above yields an estimated effect of 0.55 and a standard error of 0.38, which is statistically insignificant (p=0.15). The results are also not statistically significant ($\hat{\beta} = 0.07, SE = 0.06,\ \mathrm{p} = 0.24$) if the outcome variable is replaced with ln(Condom Purchases + 1).

[38] In particular, a negative binomial regression for # Sex Acts in Past Week yields the same estimated mean marginal effect as OLS to the third decimal place.

[39] As noted above, I use the adjusted versions of the belief variables, which removing time-varying trends in beliefs. All the results are robust to using the original belief variables as well.

and $\hat{\delta}$ is an estimate of $\frac{dy}{dx}$, the marginal effect of risk beliefs on risky sex. The results of these regressions are shown in Panel A of Table 8, and discussed below. However, for these estimates to be consistent, x_{i1} must be independent of the error term. This is unlikely to be true. One reason it may fail is that individuals may form their risk beliefs based in part on sexual experience, and sexual experience is highly autocorrelated. Another, noted by Oster (2012), is that the subjective risk will probably have some association with the actual prevalence HIV – but that prevalence is itself the outcome of local sexual behavior.

I therefore estimate $\hat{\delta}$ via two-stage least squares, using T as an instrument for x. T is plausibly excludable. The evidence presented in Section 3 demonstrates that the experiment was successfully randomized, so that membership in the treatment group should have no association with sexual behavior other than through the information treatment. Furthermore, the information treatment is very unlikely to affect sexual behavior through any channel other than individuals' risk beliefs: it does not contain any guidance or information about sex. The instrument also easily satisfies the relevance condition. The F-statistic on T in the first-stage regressions is roughly 220 for all specifications except condom use, where it is 100 (the condom use regression has a smaller sample because it is conditional on having had sex).[40] This allows me to estimate two-stage regressions as follows:

$$x_{i1} = \alpha + \beta_{ANCOVA} T_{i1} + \gamma y_{i0} + \rho x_{i0} + \eta Z_{i0} + e_i$$
$$y_{i1} = \alpha + \delta \hat{x}_{i1} + \gamma y_{i0} + \rho x_{i0} + \eta Z_{i0} + e_i$$

x_{i0} is included as a control in the first stage in order to improve efficiency and reduce bias, for the same reason discussed in Section 4.1 above.

The 2SLS estimates are shown in Panel B of Table 8, with OLS results shown in Panel A for comparison. The OLS results have a uniform positive bias relative to 2SLS, confirming that OLS is not consistent in this context. The estimated marginal effects, like the ITT effects, are small but statistically significant for the sex diary outcomes, and not distinguishable from zero for the other outcomes. Panel C divides the point estimates and endpoints of the 95% confidence interval by the sample mean as one measure of the semi-elasticity. Alternatively, I estimate Column 2 in logs instead of levels (adding 1 to the underlying variable to account for zeroes)

[40] It is not possible to conduct a formal test for weak instruments unless the number of excluded instruments is at least two more than the number of endogenous regressors. However, the informal "rule of thumb" generally used in applied econometrics is an F-statistic of at least 10; by this standard, my instrument easily passes.

yielding an estimated elasticity of -0.5 with a standard error of 0.2.

Both the population-average ITT effects and marginal effects fit a model of self-protection, which is consistent with the existing literature. However, the specifications in Tables 7 and 9 impose common effects across all respondents, and hence across all levels of risk beliefs. To explore the importance of this restriction, I explore heterogeneity in ITT and marginal effects by baseline covariates, with a focus on baseline risk beliefs.

4.6. Heterogeneous impacts of risk information on sexual risk-taking

The key prediction of the rational fatalism model is that responses to risks will be heterogeneous by individuals' baseline characteristics. Specifically, it predicts that the magnitude and sign of the comparative static will vary by baseline beliefs about risks. In earlier work (Kerwin 2012) I discuss the potential of nonparametric methods for studying the variation in question. However, the field experiment I analyze in this paper involves only one treatment and one control group, and hence is summarized in the data by a single binary variable. This limits my analysis to the estimation of linear local average treatment effects.[41]

In lieu of fitting a nonparametric model to the experimental data, I focus heterogeneous responses to the information treatment by baseline covariates. In general, to examine heterogeneity in intent-to-treat effects by a set of J baseline covariates $w_1, ..., w_J$ I follow de Mel et al. (2008) in estimating specifications of the form:

$$y_{i1} = \alpha + \beta^{T}_{ANCOVA} T_{i1} + \sum_{j=1}^{J} \left[\beta^{Tw_j}_{ANCOVA} T_{i1} w_{ji0} + \delta_j w_{ji0} \right] + \gamma y_{i0} + \eta Z_{i0} + e_i$$

where w_{ji0} is individual i's value of w_j measured at baseline, y_{i1} and y_{i0} are the outcome measured at endline and baseline respectively, and Z_{i0} is a vector of other controls[42] measured at baseline. My primary focus is on heterogeneity by baseline risk beliefs (x_{i0}). I also examine other potential sources of heterogeneity in responses, such as gender, baseline sexual activity,

[41] In principle, an experiment could have included multiple study arms involving revealing different information (such as the local prevalence of HIV). However, the maximum number of different information treatments would be quite small unless some of them involved giving inaccurate information, and the logistical difficulty of providing a large number of information treatments would likely cause problems with compliance failures and information spillovers. Hence any feasible experimental study of risk beliefs and risky sex could not use high-order polynomials or semiparametrics due to an inability to identify enough parameters.

[42] All specifications include controls for the original sampling strata (categorical distance from nearest trading center interacted with gender).

and the previous HIV exposures.

The results of these heterogeneous treatment effects analyses for the total number of sex acts in the past week are presented in Table 9. Responses to the information treatment are strongly heterogeneous by baseline risk beliefs (Column 1). People with baseline risk beliefs of 0% respond to the information treatment by increasing their sex acts per week by 0.7. For people with baseline beliefs of 100%, the response is lower by 1.2 sex acts, meaning that sexual activity *declines* by 0.5 acts per week. I can reject that responses for people with high risk beliefs are the same as for those with low beliefs at the 1% level; the negative response for people with the highest risk beliefs is statistically significant at the 10% level.

Another dimension by which responses to the treatment might plausibly vary is the gender of the respondent, which I explore in Column 2. It is theoretically ambiguous whether differential responses by gender should be expected in my sample. Malawian women commonly have less bargaining power in sexual relationships than men. However, most of my sample comprises matrilocal villages, which grant women more power to divorce their husbands and hence may increase bargaining power within relationships as well (Schatz, 2005). The results show that males respond less positively to the treatment, but this difference is not statistically significant. In Columns 4 through 6 I look for heterogeneous responses by baseline sexual activity, perceived previous exposure to HIV,[43] and whether the respondent believes he or she may currently be HIV positive.[44] There is also no statistically significant heterogeneity by any of these factors. Moreover, the results for baseline risk beliefs are also robust to including three-way interactions with gender, as well as the other variables in Table 9.

The specification in Table 9 assumes that the heterogeneity in treatment effects is linear in form. While this is not a concern for binary w_j such as gender, it is a more substantive restriction for continuous variables like baseline beliefs. As an alternative, I break the baseline risk belief variable into brackets: <10%,10%-20%,…,90%+. I then run the same regression specification but using binary indicators for each of these brackets as the w_js of interest. The

[43] Perceived previous exposure to HIV is an indicator that is coded to 1 if the respondent believes any of their past sex partners was HIV-positive and zero if they do not. This ignores the possibility that a condom was used for the sex acts with an HIV-positive partner, but given the low rates of condom use in this population that should not affect the results appreciably.

[44] The perceived HIV status variable is an indicator that collapses a Likert scale question in which respondents report how likely they think it is that they are HIV-positive now on a scale from "No Likelihood" up to "High Likelihood". "No Likelihood" is coded as a zero, while any other response is coded as a one. "Don't Know" is coded as a missing value.

results of this regression are plotted in Panel A of Figure 9, with bootstrapped 90% confidence intervals. I also run a similar regression for the first stage, with results shown in Panel A of Figure 10.

As an alternative way of capturing heterogeneity in treatment effects flexibly, I estimate semiparametric regressions of dy/dT by baseline risk beliefs for the treatment and control groups and take their difference,

$$y_{i1} = \beta_0^T + f^T(w_{i0}) + \gamma^T y_{i0} + \eta^T Z_{i0} + \varepsilon_i \text{ if Treatment = 1}$$
$$y_{i1} = \beta^C + f^C(w_{i0}) + \gamma^C y_{i0} + \eta^C Z_{i0} + \nu_i \text{ if Treatment = 0}$$

The semiparametric regressions are implemented using the Verardi and Debarsy (2012) implementation of the Robinson (1988) double residual estimator for partially linear regressions.[45] The basic logic of the Robinson estimator is as follows: consider the regression function for the control group. If we take its conditional expectation given w_{i0}, and subtract that from the original equation, the $f(w_{i0})$ component drops out and we have

$$y_{i1} - \mathbb{E}[y_{i1}|w_{i0}] = \gamma^C(y_{i0} - \mathbb{E}[y_{i0}|w_{i0}]) + \eta^T(Z_{i0} - \mathbb{E}[Z_{i0}|w_{i0}]) + \nu_i$$

.The conditional expectations of y_{i1} given w_{i0} and so forth, are estimated by separate non-parametric regressions for each variable. Then these estimates are plugged in to the equation above, which is estimated by OLS. Finally, the parametric component of y_{i1} is removed using the estimates of γ^C and η^C, allowing the function $f^C(w_{i0})$ to be estimated non-parametrically.

Having constructed $f^T(w_{i0})$ and $f^C(w_{i0})$, I then apply a procedure similar to Bennear et al. (2011), who use the difference between simple non-parametric regressions of y_{i1} on w_{i0} for the treatment and control groups as a measure of heterogeneity in treatment effects. I construct the treatment effect for any specific value of w_i, $\tau_y(w_{i0})$, as $\tau_y(w_{i0}) = f^T(w_{i0}) - f^C(w_{i0})$. I apply this approach to heterogeneity in treatment effects by baseline risk beliefs, estimating a function $\tau_y(x_{i0})$. I also conduct a parallel procedure for endline beliefs x_{i1}, and construct a function $\tau_x(x_{i0})$. I then construct confidence intervals via a clustered bootstrap with 1000

[45] The alternative Yatchew (1997) differencing estimator is problematic in this context because the belief variable that is treated nonparametrically has tied values. No explicit guidance on handling this issue is available from the original paper or from the Lokshin (2006) article on implementing the estimator in Stata. One reasonable approach is to randomly re-sort the tied observations and use the average of the estimates thus generated to partial out the parametric component of y. When this approach is implemented, the results are substantively identical to what is found using the Robinson double residual method.

repetitions.[46] Panel B of Figure 9 shows the results of this semiparametric regression for the ITT effect, and Panel B of Figure 10 shows the results for the first stage.

The results shown in Figure 9 are consistent with the linear approximation to the heterogeneity in intent-to-treat effects: the treatment effect starts positive, and then becomes negative for people with extremely high baseline risk beliefs. For people with the highest baseline beliefs, I can reject the null that the treatment effect is ≥ 0 at the 5% level (p=0.27 for the specification in Panel A). The first-stage results in Figure 10 show that the change in risk beliefs is largest for people with the highest beliefs, and drops fairly steadily for as baseline beliefs fall.[47] This pattern is reasonable, since people with the highest risk beliefs should update their priors further than people with lower beliefs. It also suggests that the marginal effect of beliefs on behavior will be even more strongly heterogeneous than the ITT effect of the treatment.

4.7. Heterogeneous marginal effects of risk beliefs on behavior

My theoretical model predicts not just heterogeneity in treatment effects but also heterogeneity in the effect of risk beliefs x on sexual behavior y. In particular, it implies that the marginal effect of x on y will be initially negative, and then positive for sufficiently high initial risk beliefs x_{i0}. Hence it is important to look at heterogeneity in the instrumental variables estimate of the effect of x on y.

Consider some consistent instrumental variables estimator of the marginal effect dy/dx for the entire sample, $\hat{\alpha}_{IV}$. Define subgroup k of the sample as those individuals with some baseline covariate $w_{i0} = w^k$. It is possible to construct an estimator of the group k-specific marginal effect $\hat{\delta}^k_{IV} = \hat{\delta}^k_{IV}(w^k)$, which will in general be a function of w^k. Since T and w_{i0} are independent, the treatment remains a valid instrument for this subsample. Selection on right-hand

[46] Because of the large number of non-parametric regressions to be run for each bootstrap iteration, and the fact that the optimal bandwidth choice might vary based on the bootstrap sample in question, I rely on data-driven approaches for selecting the bandwidths for these regressions. In particular, I use the rule-of-thumb bandwidth of Fan and Gijbels (1996), which minimizes the conditional weighted mean integrated squared error. I also consistently use a Gaussian kernel for all non-parametric regressions.

[47] Near the low end of the scale the estimated dx/dT is larger in magnitude than actual beliefs x_0. This happens because dx/dT is estimated off of endline beliefs, and beliefs rose by less in the treatment group than they did in the control group. For example, for people with baseline beliefs below 0.1 the average endline belief was 0.18 in the control group and 0.10 in the treatment group. This suggests that there is some degree of mean-reverting measurement error in the belief variables, so that some of the extreme values represent people with more moderate beliefs whose data was mismeasured.

side variables likewise does not affect the consistency of an estimator, so any valid instrumental variables estimator for the whole sample will be valid for this subsample. While I could rely on 2SLS estimation, in general I will want to estimate the relationships semiparametrically, so I instead use the indirect least squares (ILS) estimator. I estimate the following separate regressions:

$$x_{i1} = \alpha^x + \beta^x T_i + \gamma^x x_{i0} + \delta^x Z_i + e_i \text{ for } w_{i0} = w^k$$
$$y_{i1} = \alpha^y + \beta^y T_i + \gamma^y y_{i0} + \delta^y Z_i + v_i \text{ for } w_{i0} = w^k$$

where x_{i0} and y_{i0} are the baseline values of x and y respectively and Z is a vector of other controls. These are ANCOVA regressions for the reasons outlined above. I then construct

$$\hat{\delta}_{ILS,j}(w^k) = \frac{\hat{\beta}^y(w^k)}{\hat{\beta}^x(w^k)} \xrightarrow{p} \frac{\frac{dy}{dT}(w^k)}{\frac{dx}{dT}(w^k)} = \frac{dy}{dx}(w^k)$$

Where convergence in probability comes from Slutsky's theorem, or by appealing to the derivation of ILS. In practice, I will recover the treatment effects $\hat{\beta}^y(w^k)$ and $\hat{\beta}^x(w^k)$ using $\tau^y(w^k)$ and $\tau^y(w^k)$ as described above. While it is possible to construct analytic standard errors for ILS, I rely instead on bootstrapped confidence intervals since my preferred underlying estimator is already semiparametric and has standard errors without a know analytical form.

An alternative approach along the same lines is to estimate $\hat{\beta}^y(w^k)$ and $\hat{\beta}^x(w^k)$ parametrically, using indicators for different levels of w_{i0}. I break w_{i0} into C categories and construct indicators w_{i0}^c for membership in each category. I then run a saturated regression that interacts categorical indicators for the categories of w_{i0} with the treatment indicator as well as all other controls:

$$y_{i1} = \sum_{c=1}^{C} T_{i1}\left[\beta_{ANCOVA}^{Tw^c} w_{i0}^c + \delta^{w^c} w_{i0}^c + \gamma^{w^c} w_{i0}^c y_{i0} + \eta^{w^c} w_{i0}^c Z_{i0}\right] + e_i$$

The point estimates of $\hat{\beta}_{ANCOVA}^{Tw^c}$ from this specification will be identical to running separate regressions for each category. However, the estimated standard errors may be affected by the pooled regression. Instead of using a weighting approach to adjust the residuals, I construct bootstrap confidence intervals using stratified subsamples.

The results of the parametric specification are shown in Panel A of Figure 11, while the semiparametric results are in Panel B. These marginal effects are consistent with the theoretical framework from Section 2, in which the relationship between risk beliefs and risky sex has an

overall U-shape: the slope is initially negative and then becomes positive for people with sufficiently high risk beliefs. I reject the null that marginal effects are ≥ 0 for the highest risk belief category at the 5% level (p=0.21 for the parametric specification). Note that although this evidence suggests a U-shaped relationship, I am unable to recover the underlying function: I can estimate heterogeneity in the marginal effect of endline risk beliefs on risky sex only by *baseline* risk beliefs, not by endline beliefs.

Relationship between overall and covariate-specific LATEs

The w^k-specific LATEs estimated by the above procedure form the components of the overall LATE for the entire sample, but the overall LATE is a weighted average of these components, not a simple mean. In this subsection I show that their weights in forming the overall LATE are given by the share of the data with each w^k times the degree of compliance with the instrument (the extent to which the instrument shifts x) for each w^k.

For the sake of exposition, begin with the Wald IV estimator of the local average treatment effect for the whole population, given by

$$\hat{\delta}_{Wald} = \frac{\mathbb{E}[y_i|T_i = 1] - \mathbb{E}[y_i|T_i = 0]}{\mathbb{E}[x_i|T_i = 1] - \mathbb{E}[x_i|T_i = 0]}$$

This is asymptotically equal to the ILS and 2SLS estimators because all three are consistent. By the Law of Iterated Expectations, $\mathbb{E}[y_i|T_i = 1] = \mathbb{E}[\mathbb{E}[y_i|T_i = 1]]$. Assume the baseline covariate w_{i0}, is discrete (or measured discretely due to the data collection process) with values $w_1, ..., w_K$. Using the Law of Total Expectation, we can rewrite $\mathbb{E}[\mathbb{E}[y_i|T_i = 1]]$ as $\sum_{k=1}^{K} \mathbb{E}[\mathbb{E}[y_i|T_i = 1]|w_{i0} = w^k]\mathbb{P}(w^k)$. Then we have that

$$\begin{aligned}\hat{\delta}_{Wald} &= \frac{\sum_{j=1}^{m} \mathbb{E}[\mathbb{E}[y_i|T_i = 1]|w_{i0} = w^k]\mathbb{P}(w^k) - \sum_{k=1}^{K} \mathbb{E}[\mathbb{E}[y_i|T_i = 0]|w_{i0} = w^k]\mathbb{P}(w^k)}{\sum_{k=1}^{K} \mathbb{E}[\mathbb{E}[x_i|T_i = 1]|w_{i0} = w^k]\mathbb{P}(w^k) - \sum_{k=1}^{K} \mathbb{E}[\mathbb{E}[x_i|T_i = o]|w_{i0} = w^k]\mathbb{P}(w^k)} \\ &= \frac{\sum_{k=1}^{K} \mathbb{E}[\mathbb{E}[y_i|T_i = 1] - \mathbb{E}[y_i|T_i = 0]|w_{i0} = w^k]\mathbb{P}(w^k)}{\sum_{k=1}^{K} \mathbb{E}[\mathbb{E}[x_i|T_i = 1] - \mathbb{E}[x_i|T_i = 0]|w_{i0} = w^k]\mathbb{P}(w^k)}\end{aligned}$$

Let $\hat{\beta}^y(w^k)$ be a consistent estimate of the effect of the treatment on y given $w_{i0} = w^k$, $\mathbb{E}[\mathbb{E}[y_i|T_i = 1] - \mathbb{E}[y_i|T_i = 0]|w_{i0} = w^k]$. Then the Wald estimator can be rewritten as $\frac{\sum_{j=1}^{m} \hat{\beta}^y(w^k)\mathbb{P}(w^k)}{\sum_{j=1}^{m} \hat{\beta}^x(w^k)\mathbb{P}(w^k)}$. The ILS estimator for a w^k-specific slope is $\hat{\alpha}_{ILS,j}(w^k) = \frac{\hat{\beta}^y(w^k)}{\hat{\beta}^x(w^k)}$,

So we can rewrite the Wald estimator as:

$$\frac{\sum_{j=1}^{m} \hat{\alpha}_{ILS}(w^k)\hat{\beta}^x(w^k)\mathbb{P}(w^k)}{\sum_{j=1}^{m} \hat{\beta}^x(w^k)\mathbb{P}(w^k)} = \sum_{j=1}^{m} \hat{\alpha}_{ILS}(w^k)\frac{\hat{\beta}^x(w^k)}{\hat{\beta}^x}\frac{n_j}{N}$$

where n_j is the number of observations with $w_{i0} = w^k$ and N is the total number of observations in the dataset. Let $\theta_j = \frac{\hat{\beta}^x(w^k)}{\hat{\beta}^x} n_j$. Then we have $\hat{\delta}_{Wald} = \frac{\sum_{j=1}^{m} \hat{\delta}_{Wald}(w^k)\theta_j}{N}$. The overall estimate of the slope of y with respect to x is the weighted average of the w^k-specific slope estimates.

The weights θ_j have two components. The first part is the number of observations with $w_{i0} = w^k$. This is multiplied by the second part: the ratio of the impact of the treatment on x at $w_{i0} = w^k$ to the overall treatment effect, which is a continuous measure of compliance with the categorical instrument T. Observations where the treatment shifts x more have greater weight in determining the overall mean marginal effect estimate – in other words, the overall LATE is the compliance-weighted average of the baseline covariate-specific LATEs.

Running this computation with the estimates from Panel B of Figure 11 gives a weighted average of -1.25, as compared with a 2SLS estimate of -1.17 from Column 2 of Table 8. The small difference can be attributed to the fact that I control linearly for baseline risk beliefs in Table 8, which is not a sensible approach when I am breaking up the same by baseline beliefs as well. The regressions computations underlying the ILS estimates used for Panel B also do not throw out values of y when x is missing (or vice versa) meaning they are estimated on a slightly larger sample.

5. Discussion

The heterogeneity in the marginal effect of risk beliefs on risky sex raises important questions about the whether the mean marginal effect has any useful epidemiological meaning. At the same time, my focus on this specific aspect of heterogeneity also raises potential questions about multiple comparisons and specification choice. Section 5.1 explores the mechanisms behind the fatalistic responses found in this population. In Section 5.1, I discuss what the results from Section 4 mean for the effect of scared-straight style policies on the prevalence of HIV, while in Section 5.2 I focus specifically on the direct effect of this information treatment on HIV infections. Section 5.3 addresses several potential limitations of this study, in particular problems

of multiple comparisons.

5.1. Mechanisms for Fatalistic Responses

The theoretical framework in Section 2 predicts that fatalistic responses to risks in two different situations. First, people may have an accumulated stock of past risks they have taken whose outcome has not yet been realized. Second, they may not have perfect control over their future risky behavior: condoms may break, they may be tempted into mistakes, and so forth. The results in Section 4 cannot distinguish between these two mechanisms.

One way of evaluating which mechanism is at work is to examine whether the effect of the information treatment varied based on people's beliefs about their current HIV status. If the first mechanism alone is driving the fatalism measured in our sample, then people's responses to the information treatment should be negative if (and only if) they believe they are currently HIV-positive. There is no evidence of this pattern in my sample: Column 6 of Table 9 shows that there is no statistically-significant difference in the treatment effect by people's baseline beliefs about their HIV status.[48] I also find no evidence of effects on respondents' endline beliefs about their HIV status. Figure 12 replicates the analysis in Panel B of Figure 9 but using an indicator for whether people think they may be HIV-positive as the outcome. There are no statistically-significant effects on this belief variable at any level of baseline risk belief.

In contrast, the treatment does significantly reduce people's likelihood of reporting that they may contract HIV in the *future* – but only for the portion of the sample that responds fatalistically to the information treatment. Figure 13 mirrors Figure 12, but using an indicator for whether people think they may contract HIV in the future.[49] There is a statistically-significant reduction in this variable for people with the highest risk beliefs: people with these extremely high beliefs are around 20 percentage points less likely to think they will contract HIV in the future if they receive the information treatment, as compared to the control group. Furthermore, the pattern looks similar to that for the response to the information treatment.

One reason that fatalism appears to operate only through future HIV infections in this

[48] See Section 4.6 for a description of how this variable is defined.

[49] The perceived HIV status variable is an indicator that collapses a Likert scale question in which respondents report how likely they think it is that they are HIV-positive now on a scale from "No Likelihood" up to "High Likelihood". "No Likelihood" is coded as a zero, while any other response is coded as a one. "Don't Know" is coded as a missing value. The results in Figure 13 are unaffected by the exclusion of people who say they already know they have HIV in response to this question.

population may be the availability of HIV testing. If people have easy access to HIV tests, they would only falsely assume that past exposures had made them HIV-positive if they were in the "window period" before blood testing can detect the virus. It does appear that HIV testing is readily and cheaply available in the local area. 95% of my sample says they know a place where they can get tested for HIV, and all but one of the people who knows of a place where HIV tests are available is able to name it. The time and pecuniary costs of testing are also perceived to be low: 70% of people believe they can walk to the testing site in one hour or less (and 92% within two hours) while 92% report (accurately) that HIV tests are free of charge. Wait times for testing are also reported to be low: 63% think testing would take less than 30 minutes and 87% think it would take less than one hour. These economic costs ignore the emotional barriers to HIV testing, which are not insignificant, but suggest that motivated individuals who are afraid they already have HIV can resolve their uncertainty through testing.

In contrast, (individual) HIV testing cannot do anything to resolve people's uncertainty about whether they are doomed to contract HIV in the future. People who perceive extremely high risks of HIV infection, and know they cannot prevent all possible exposures to the virus, will rationally conclude that they cannot avoid HIV in the future even if they receive a negative test result in the present. The results in Figure 13 suggest that people with high risk beliefs had become convinced they would contract HIV no matter what they did, and that the information treatment convinced them they were wrong. Indeed, in the control group data, this subgroup us far more likely to report that they might contract HIV in the future than the rest of the population. 70% of control-group respondents who held baseline risk beliefs of 0.8 or higher said they might get HIV in the future at endline, as opposed to 55% of people with lower risk beliefs.

The lack of an effect on this variable for the rest of the population could be understood in part as a form of risk homeostasis: people at the lower end of the belief scale had more risky sex, thus compensating for the decrease in the per-act risk of contracting HIV by increasing their number of sex acts. These results suggest that the information treatment reduced the extent of fatalistic responses in this population through reducing the chance that they believed they would contract HIV in the future, rather than through reducing their likelihood of believing they already had the virus.

5.2. Epidemiological Implications

Studies of rational epidemics have generally focused on the mean marginal effect of risks on risky behaviors as a useful summary statistic for the population as a whole. However, the non-monotonic relationship between risk beliefs and risky sex suggested by the results in Section 4.6, suggests that the mean marginal effect may not tell us anything useful about how the prevalence of the virus responds to changes in people's risk beliefs. Epidemiologists have found that aggregate HIV transmission is dominated by high-activity individuals (Koopman et al. 2005). As a result, declines in risk-taking among high-activity individuals can offset rises in risk-taking by low-activity people. Since responses to risks vary by baseline risk beliefs, the crucial question is whether people's level of sexual activity varies with those beliefs as well, and, if so, how much.

Using the same results that generated Panel B of Figure 11, I predict an estimated individual treatment effect for every person in the treatment group based on their individual baseline risk belief. Table 10 shows the shares of the treatment group, and statistics on lifetime sex partners, for people who increased how much sex they had as a result of the treatment (about 86%) and for people who reduced their sexual activity (about 14%). The last two columns of Table 10 show that members of the treatment group who decreased their risk-taking as a result of the intervention have an average of 4.2 lifetime sex partners, as opposed to 3.2 for those who took more risks; this difference is statistically significant at the 0.05 level. This suggests that the individuals who appear to be fatalistic may be more represented in the high-activity portion of the population that is crucial for determining the spread of the epidemic. The policy of attempting to scare people straight is thus backfiring for exactly the population we most want to target for risk-reduction interventions. To know whether the scared-straight approach is working – and thus whether teaching people more accurate risk information reduce new HIV infections – would require an epidemiological model of the disease that explicitly incorporates potentially non-monotonic responses to risk beliefs on the part of the agents in the model. Relying on estimated mean marginal effects cannot tell us anything useful about how the population-level prevalence of the virus responds to people's risk beliefs.

5.3. Direct Effects on HIV infections

One question that can be addressed directly is the impact of this specific information intervention on additional HIV infections. To compute an effect on HIV infections I would ideally rely on HIV test results, but these were not collected as part of the study. Instead, I use

two proxies. First, respondents' self-reported beliefs about their own, and their primary sex partner's, serostatus, which are measured on the survey. Second, data from the 2010 DHS, which measured the prevalence of HIV in Zomba District, but cannot be directly tied to my responses.

For the first proxy, I consider as serodiscordant couples in which the respondent reports no likelihood of being HIV infected him- or herself, but some likelihood for his or her partner, or vice-versa. Using the same method as in Table 10, the total change in weekly sex acts for all 77 respondents in such couples (summing over the changes for each individual respondent) is 4.39. Respondents in this group used condoms just 3.7% of the time, so this would mean 4.23 more unprotected acts per week. This would correspond to an additional 0.004 HIV infections per week total.

To use the second proxy, I first note that the 2010 DHS measured an HIV prevalence of 18% for Zomba District. The DHS data has too few instances of multiple partners within a household to estimate the degree of matching on HIV status among couples in Zomba District, so instead I assume that people pick their sex partners randomly with respect to HIV status. The heterogeneity in people's responses to the treatment will tend to work against an increase in HIV transmissions. I can therefore find an upper bound by ignoring any heterogeneity and treating the effect (0.2 sex acts per week, from Column 2 of Table 6) as constant. Under these assumptions, the 18% of the treatment group that are HIV-positive have an additional (0.20)X(0.82) = 0.16 sex acts each week with HIV-negative partners, while the 82% of my sample that is HIV-negative has an additional (0.20)X(0.18) = 0.04 sex acts each week with HIV-positive sex partners. Taking the weighted sum, and multiplying by the fraction of sex acts that are unprotected (88%), the average person in the treatment group would have an additional 0.05 sex acts per week where an HIV transmission was possible. That would mean a total of 34 such sex acts per week across the 647 people in the treatment group. This would correspond to 0.03 additional HIV transmissions per week in the absence of any other changes.

This rise in the number of HIV infections due to the information treatment does not necessarily mean that the study as a whole caused more people to contract HIV. All respondents were also sold heavily discounted condoms as part of the followup survey, and on average people in the treatment group bought 5.22. At baseline, my respondents had used about 20% of the condoms they had gotten in the past 30 days. Making the assumption that the same will hold for these condoms, that would mean an additional 199.4 condoms used in sex acts where one

partner is HIV positive over the next 30 days. This would more than compensate for the additional 144.1 unprotected acts that happened as a result of the information treatment, using the upper-bound figures constructed using the DHS. A similar pattern holds over the longer term: if I assume the information treatment's effects persist without diminishing for six months, my upper-bound estimates would predict 873 additional sex acts between an HIV-positive and HIV-negative, while if 100% of the condoms that were bought were eventually used, this would imply 997 additional protected sex acts among this group.[50] It is only if the effects of the information treatment on sexual behavior persist for longer than a six-month horizon that we would expect a net increase in HIV transmissions as a result of the study.

5.4. Potential Limitations

As discussed in Section 3.2, this paper relies almost solely on self-reported sexual behavior as a measure of sexual risk-taking. This could conceivably bias my results, but in my specific context there is no reason to believe that there would be differential social-desirability bias across study arms: the information treatment provided no direct modeling of "good" behavior nor encouragement to behave in a specific way. Furthermore, my approach has the advantage of capturing changes in behavior among high-risk individuals, which cannot be done when using STIs as outcome measures unless treatable STIs are used and individuals are treated for existing STIs at baseline.

My analyses of heterogeneous treatment effects are potentially subject to the Deaton (2009) critique that subgroup analyses can constitute ex post "fishing expeditions". However, that concern is mitigated due to the fact that my main theoretical results were already laid out in earlier work (Kerwin 2012 WP). The preliminary empirical results in that paper also focus on the same primary outcome variable (total sex acts in the past week) as well as the same measure of risk belief variable (the risk of HIV infection from a single unprotected sex act with a randomly-selected attractive person from the local community).

6. Conclusion

Empirical research on behavioral responses to health risks has traditionally assumed that responses are monotonically negative, or "self-protective", and has therefore focused on mean

[50] Although some of these condoms might substitute for others the respondent intended to purchase, this is relatively unlikely given the low rates of existing condom use in the sample.

marginal effects. I use a randomized field experiment in rural Southern Malawi to explore the validity of this assumption in the context of behavioral responses to HIV infection risks. The experiment provided the treatment group with information on the true risk of HIV transmission from unprotected sex with an infected partner, which is much lower than most respondents thought. I find that the mean marginal effect of HIV risk beliefs on sexual behavior is small but statistically significant, with an elasticity of about -0.5. I develop a method to allow for heterogeneity in marginal effects (as opposed to just intent-to-treat-effects) and find that the average marginal effect masks significant heterogeneity. The effect of risk beliefs on risky sex is positive for people who initially hold low risk beliefs, and becomes negative as initial risk beliefs become sufficiently high.

This heterogeneity is consistent with a model of rationally fatalistic behavior in which changes in beliefs risk affect agent's choices not only via the risky sex acts being chosen at present, but also through a stock of previous – or unavoidable future – risky sex acts. A rise in per-act risks increases the marginal cost of more risky sex due to the first channel, but also raises the chance that HIV is simply unavoidable, which lowers the marginal cost of additional risk-taking. I show that for this population, fatalistic responses appear to be driven by people who believe that they are doomed to contract HIV in the future (for example because of condom breaks) rather than those who think they already have the virus. This may be because of the high availability of HIV testing in the local area.

My results imply that the use of mean marginal effects as a way to study the response of health behaviors to health risks may be misleading. In the case of HIV in particular, the impact of increased risk beliefs on the prevalence of the virus will depend on the importance of the self-protective and fatalistic subgroups for the epidemic. My data suggests that the fatalistic group is initially higher-activity and thus more significant in driving the epidemic.

Further research is needed on explicitly incorporating agents' perceived risk of HIV infection into rational epidemic models of HIV, rather than just assuming agents understand the true prevalence and transmission rate of the virus. Such models should also allow for responses to perceived risks to be heterogeneous by the level of the perceived risk, rather than imposing that they are the same across the whole population. The formation of people's risk beliefs is another important area for study. While anecdotal evidence suggests that people learn about HIV in school, the exact process by which many people arrive at gross overestimates of the

prevalence and transmission rate of the virus is still unknown. Given that overestimating HIV risks seems to scare people to death, rather than scaring them straight, getting at the source overestimates may be crucial for understanding the continued spread of the African HIV epidemic.

References

Ahituv, A., Hotz, V. J., & Philipson, T. (1996). The responsiveness of the demand for condoms to the local prevalence of AIDS. *The Journal of Human Resources*, *31*(4), 869–897.

Attanasio, O. (2009). Expectations and perceptions in developing countries: Their measurement and their use. *American Economic Review*, *99*(2), 87–92.

Baird, S. J., Garfein, R. S., McIntosh, C. T., & Özler, B. (2012). Effect of a cash transfer programme for schooling on prevalence of HIV and herpes simplex type 2 in Malawi: a cluster randomized trial. *The Lancet*, 1–10. doi:10.1016/S0140-6736(11)61709-1

Belli, R. F., Shay, W. L., & Stafford, F. P. (2001). Event history calendars and question list surveys: A direct comparison of interviewing methods. *Public Opinion Quarterly*, *65*(1), 45–74. doi:10.1086/320037

Bennear, L. S., Tarozzi, A., Pfaff, A., Soumya, H. B., Ahmed, K. M., & van Geen, A. (2011). *Bright lines, risk beliefs, and risk avoidance: Evidence from a randomized intervention in Bangladesh* (SSRN Scholarly Paper No. ID 1695115). Rochester, NY: Social Science Research Network.

Bruhn, M., & McKenzie, D. (2009). In pursuit of balance: Randomization in practice in development field experiments. *American Economic Journal: Applied Economics*, *1*(4), 200–232.

Card, D., & Giuliano, L. (2013). *Does gifted education work? For whom?* (Working Paper). University of California, Berkeley.

Chinkhumba, J., Godlonton, S., & Thornton, R. L. (In Press). Demand for medical male circumcision. *American Economic Journal: Applied Economics*.

De Walque, D., Dow, W. H., Nathan, R., Abdul, R., Abilahi, F., Gong, E., … Krishnan, S. (2012). Incentivising safe sex: a randomised trial of conditional cash transfers for HIV and sexually transmitted infection prevention in rural Tanzania. *BMJ Open*, *2*(1).

Deaton, A. S. (2009, January). *Instruments of development: Randomization in the tropics, and the search for the elusive keys to economic development*. Lecture presented at the British Academy.

Delavande, A., & Kohler, H. P. (2009). Subjective expectations in the context of HIV/AIDS in Malawi. *Demographic Research*, *20*(31), 817–875.

Delavande, A., & Kohler, H. P. (2011). *HIV/AIDS-related expectations and risky sexual behavior in Malawi* (Working Paper).

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.

Fan, J., & Gijbels, I. (1996). *Local polynomial modelling and its applications*. London: Chapman & Hall.

Fishbein, M., & Pequegnat, W. (2000). Evaluating AIDS prevention interventions using behavioral and biological outcome measures. *Sexually Transmitted Diseases*, *27*(2), 101–110.

Frison, L., & Pocock, S. J. (1992). Repeated measures in clinical trials: analysis using mean summary statistics and its implications for design. *Statistics in Medicine*, *11*(13), 1685–1704.

Galvin, S. R., & Cohen, M. S. (2004). The role of sexually transmitted diseases in HIV transmission. *Nature Reviews Microbiology*, *2*(1), 33–42. doi:10.1038/nrmicro794

Godlonton, S., Munthali, A., & Thornton, R. (2012). *Behavioral response to information? Circumcision, information, and HIV prevention* (Working Paper). Ann Arbor, MI: University of Michigan.

Gong, E. (2010). *HIV testing & risky sexual behavior* (Working Paper).

HIV Reinfection - Positive Prevention - Reinfection. (2007, July 17). *About.com AIDS / HIV*. Retrieved March 25, 2014, from http://aids.about.com/cs/safesex/a/reinfection.htm

Hollingsworth, T. D., Anderson, R. M., & Fraser, C. (2008). HIV-1 transmission, by stage of infection. *Journal of Infectious Diseases*, *198*(5), 687–693. doi:10.1086/590501

Holmes, K. K., Levine, R., & Weaver, M. (2004). Effectiveness of condoms in preventing sexually transmitted infections. *Bulletin of the World Health Organization*, *82*(6), 454–461. doi:10.1590/S0042-96862004000600012

Hudomiet, P., Kezdi, G., & Willis, R. J. (n.d.). Stock market crash and expectations of American households. *Journal of Applied Econometrics*.

Juhn, C., Kalemli-Ozcan, S., & Turan, B. (2009). *HIV and fertility in Africa: First evidence from population based surveys* (Working Paper No. Discussion Paper No. 4473). Institute for the Study of Labor.

Kaler, A. (2003). "My girlfriends could fill a yanu-yanu bus": Rural Malawian men's claims about their own serostatus. *Demographic Research*, *Special Collection*(1).

Kerwin, J. T. (2012). *"Rational fatalism": Non-monotonic choices in response to risks* (Working Paper). Ann Arbor: University of Michigan.

Kerwin, J. T., Foley, S. M., Thornton, R. L., Basinga, P., & Chinkhumba, J. (2011). Missing safer sex strategies in HIV Prevention: A call for further research. *African Population Studies*, *25*(2), 286–297.

Kerwin, J. T., Thornton, R. L., Foley, S. M., Chinkhumba, J., & Chibwana, A. (2011). *Situational analysis of sexual behaviors and alternative safer sex strategies in-depth interview dataset*. University of Michigan and University of Malawi College of Medicine.

Kling, J. R., Liebman, J. B., & Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, *75*(1), 83–119. doi:10.1111/j.1468-0262.2007.00733.x

Kremer, M. (1996). Integrating behavioral choice into epidemiological models of AIDS. *The Quarterly Journal of Economics*, *111*(2), 549–573. doi:10.2307/2946687

Lafferty, W. E., Coombs, R. W., Benedetti, J., Critchlow, C., & Corey, L. (1987). Recurrences after oral and genital herpes simplex virus infection. *New England Journal of Medicine*, *316*(23), 1444–1449. doi:10.1056/NEJM198706043162304

Lillard, L., & Willis, R. (2001). *Cognition and wealth: The importance of probabilistic thinking* (Working Paper No. WP 2001-007) (pp. 48109–1220). University of Michigan Retirement Research Center.

Luke, N., Clark, S., Zulu, E., Westlands, N., & Eliya, K. (2009). The Relationship History Calendar: Improving the scope and quality of data on youth sexual behavior. *Brown University Typescript*.

Malawi National AIDS Commission. (2003). *National HIV/AIDS policy: A call for renewed action*. Lilongwe: Malawi National AIDS Commission (NAC).

Malawi National Statistical Office, & Opinion Research Company's Macro International (ORC-MACRO). (2010). Malawi demographic and health Survey 2010. Calverton, MD: MEASURE DHS.

McKenzie, D. (2012). Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics*, *99*(2), 210–221.

National Health Service. (2013). Quitting timeline | Why Quit | NHS SmokeFree. Retrieved July 19, 2013, from http://smokefree.nhs.uk/why-quit/timeline/

O'Donoghue, T., & Rabin, M. (2001). Risky Behavior among Youths: Some Issues from Behavioral Economics. *NBER Chapters*, 29–68.

Office of the Secretary. (1979). *The Belmont Report*.

Oster, E. (2012). HIV and sexual behavior change: Why not Africa? *Journal of Health Economics*, *31*(1), 35–49. doi:10.1016/j.jhealeco.2011.12.006

Padian, N. S., McCoy, S. I., Balkus, J. E., & Wasserheit, J. N. (2010). Weighing the gold in the gold standard: challenges in HIV prevention research. *AIDS*, *24*(5), 621–635. doi:10.1097/QAD.0b013e328337798a

Philipson, T. J., & Posner, R. A. (1993). *Private choices and public health: The AIDS epidemic in an economic perspective*. Harvard University Press Cambridge, MA.

Robinson, P. M. (1988). Root-N-consistent semiparametric regression. *Econometrica*, *56*(4), 931–954.

Schatz, E. (2005). 'Take your mat and go!': Rural Malawian women's strategies in the HIV/AIDS era. *Culture, Health & Sexuality*, *7*(5), 479–492.

Smith, D. M., Richman, D. D., & Little, S. J. (2005). HIV Superinfection. *Journal of Infectious Diseases*, 192(3), 438–444. doi:10.1086/431682

Sterck, O. (2012). Should prevention campaigns disclose the transmission rate of HIV/AIDS? Theory and application to Burundi. *SSRN eLibrary*.

Thornton, R. L. (2008). The Demand for, and Impact of, Learning HIV Status. *American Economic Review*, *98*(5), 1829.

Viscusi, W. K. (1990). Do smokers underestimate risks? *Journal of Political Economy*, *98*(6), 1253–1269.

Wawer, M. J., Gray, R. H., Sewankambo, N. K., Serwadda, D., Li, X., Laeyendecker, O., … Quinn, T. C. (2005). Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda. *The Journal of Infectious Diseases*, *191*, 1403–1409.

Weinstein, N. D., & Klein, W. M. (1996). Unrealistic optimism: Present and future. *Journal of Social and Clinical Psychology*, *15*(1), 1–8. doi:10.1521/jscp.1996.15.1.1

Wilson, N., Xiong, W., & Mattson, C. (2012). *Is sex like driving? Risk compensation Associated with Male Circumcision in Kisumu, Kenya* (Working Paper).

Yatchew, A. (1997). An elementary estimator of the partial linear model. *Economics Letters*, *57*(2), 135–143. doi:10.1016/S0165-1765(97)00218-8

Figure 1: Example of HIV Risk Messaging from a Malawian Life Skills Textbook

Case study

Read the case study below and answer the questions that follow.

Nabetha suffered from oral thrush only a month after her first sexual encounter. The infection persisted over a long period of time. The doctor recommended a HIV test. She tested positive. The results devastated her. She decided she would not suffer alone but infect as many people as possible. She worked very hard towards this end. She died three years after the diagnosis with HIV.

Figure 2: Shapes of Risk-Aggregation Functions for Low and High Values of Per-Act Risk

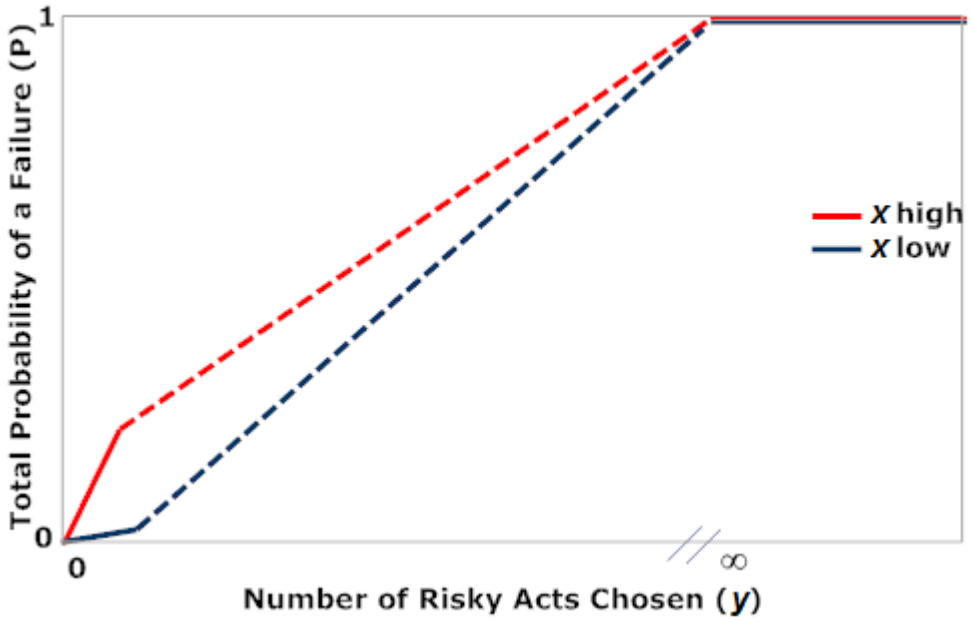


Figure 3: Histogram of sex acts reported conditional on any sex
Panel A: Single-question recall, past 30 days

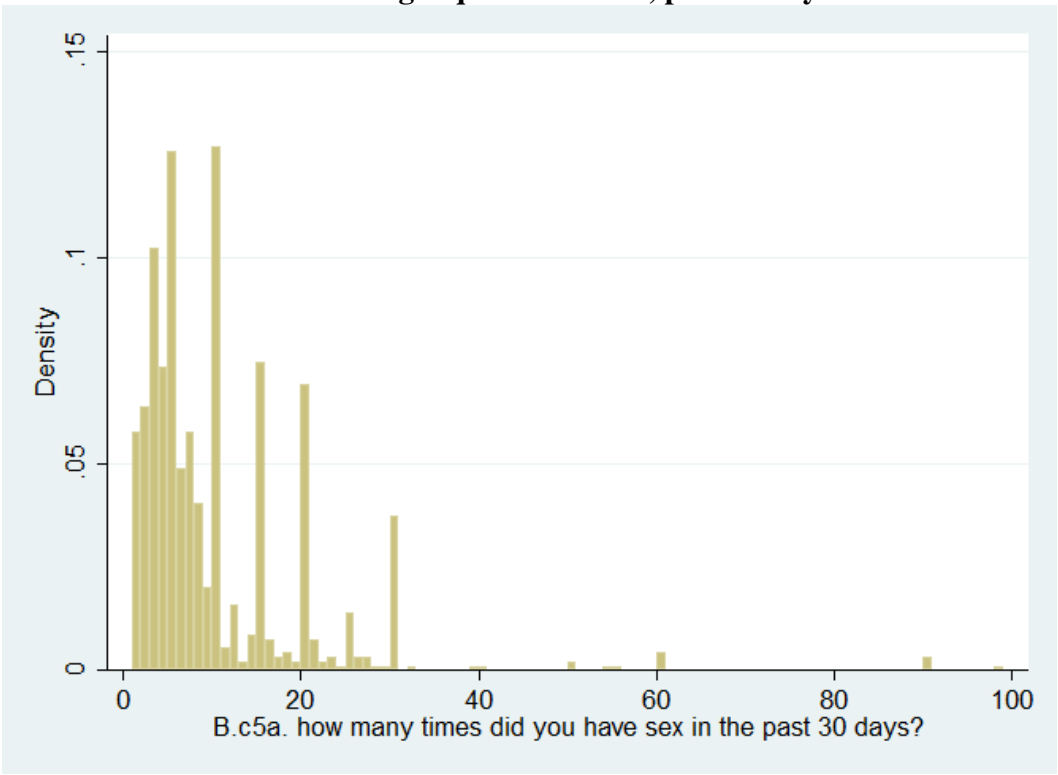


Figure 3: Histogram of sex acts reported conditional on any sex
Panel B: Retrospective sex diary, past 7 days

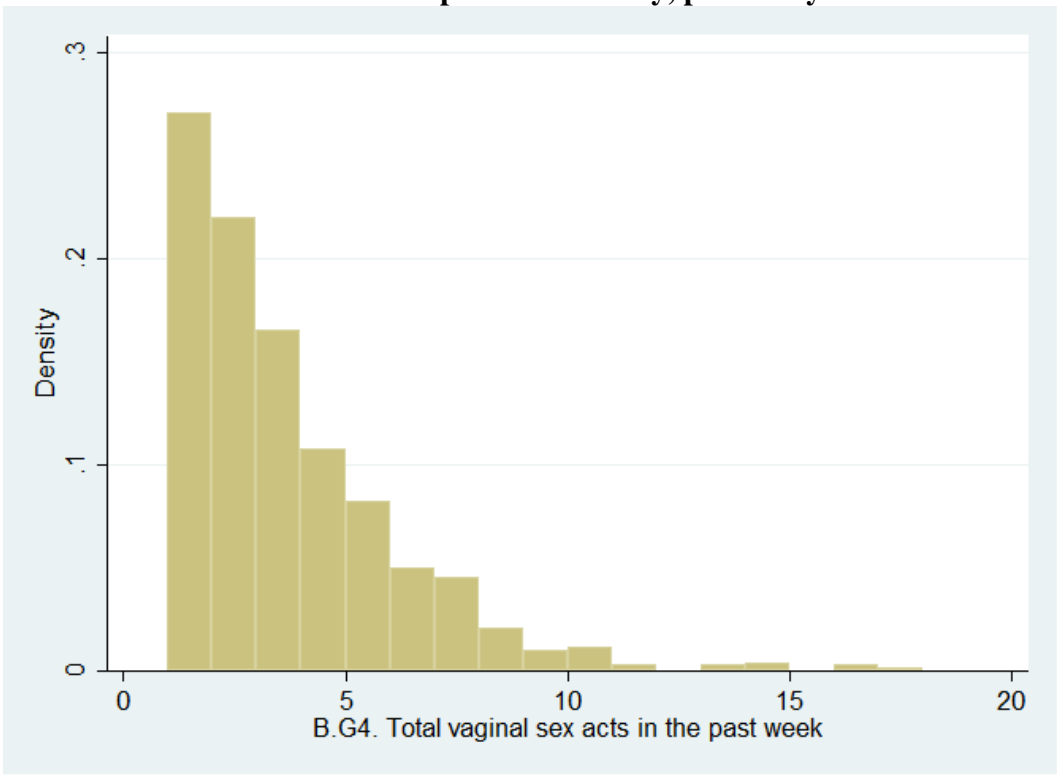


Figure 4: Example Question about Subject Risk Beliefs

E1a. If 100 men, who do **not** have HIV, each sleep with a woman who is HIV positive tonight and do **not** use a condom, how many of the
think will have HIV after the night?

Number: | # | # | # |

E1b. **If answer to E1a is 50** Do you really think that 50 of the men would get HIV, or are you just not sure?

☐ 1. I really think it's 50 ☐ 0. I'm just not sure → What is your best guess? | # | # |

Figure 5: Effect of Enumerator Knowledge on Measured Risk Beliefs
(Per-act HIV transmission rate for unprotected sex)

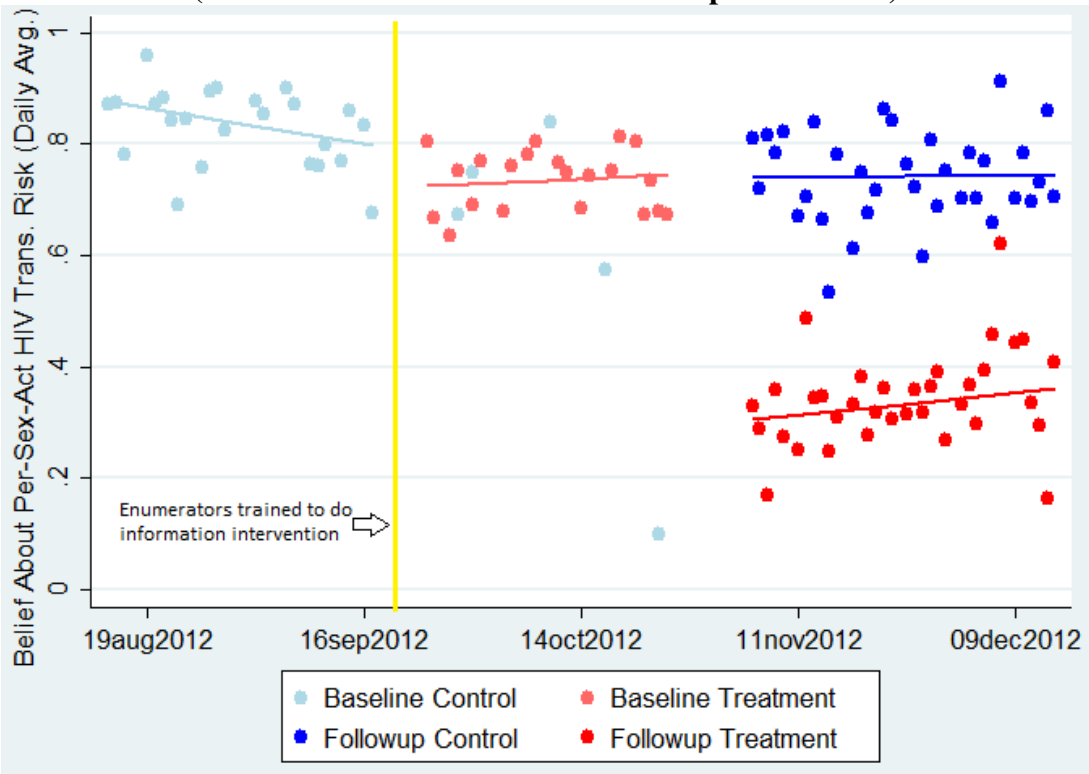


Figure 6: Histogram of Baseline Beliefs about Per-Act HIV Infection Risk from a Random Attractive Sex Partner, Unadjusted

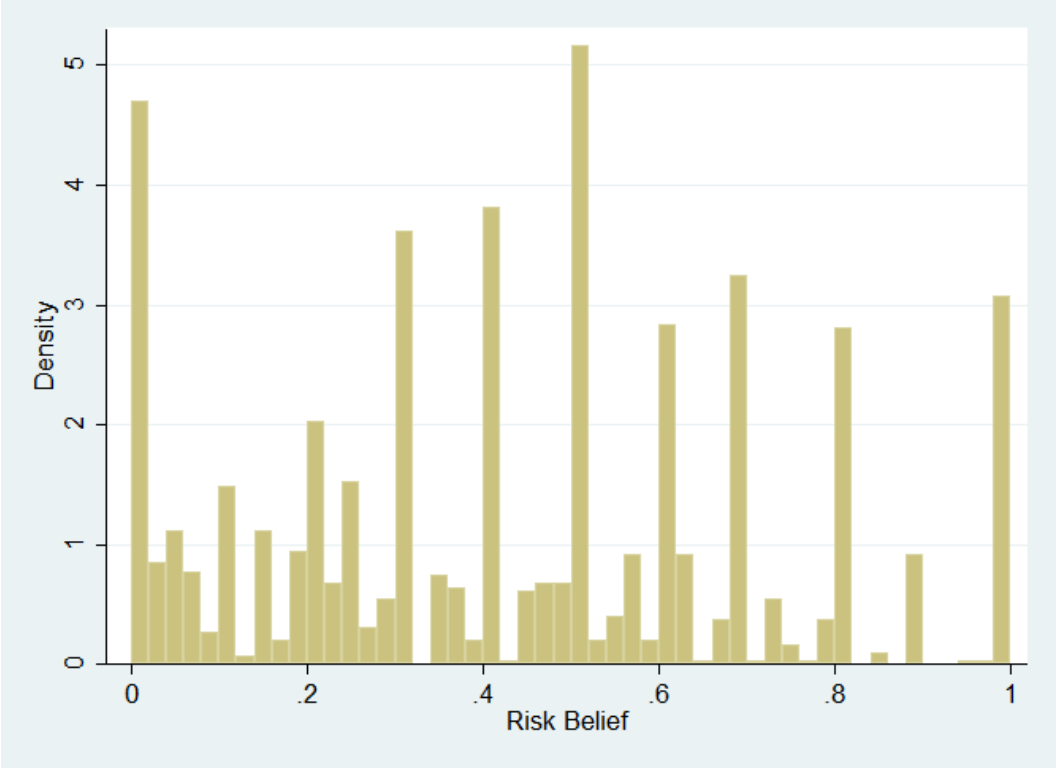


Figure 7: Kernel Density of Baseline Beliefs about Per-Act HIV Infection Risk from a Random Attractive Sex Partner, Adjusted for Non-Constant Linear Trend

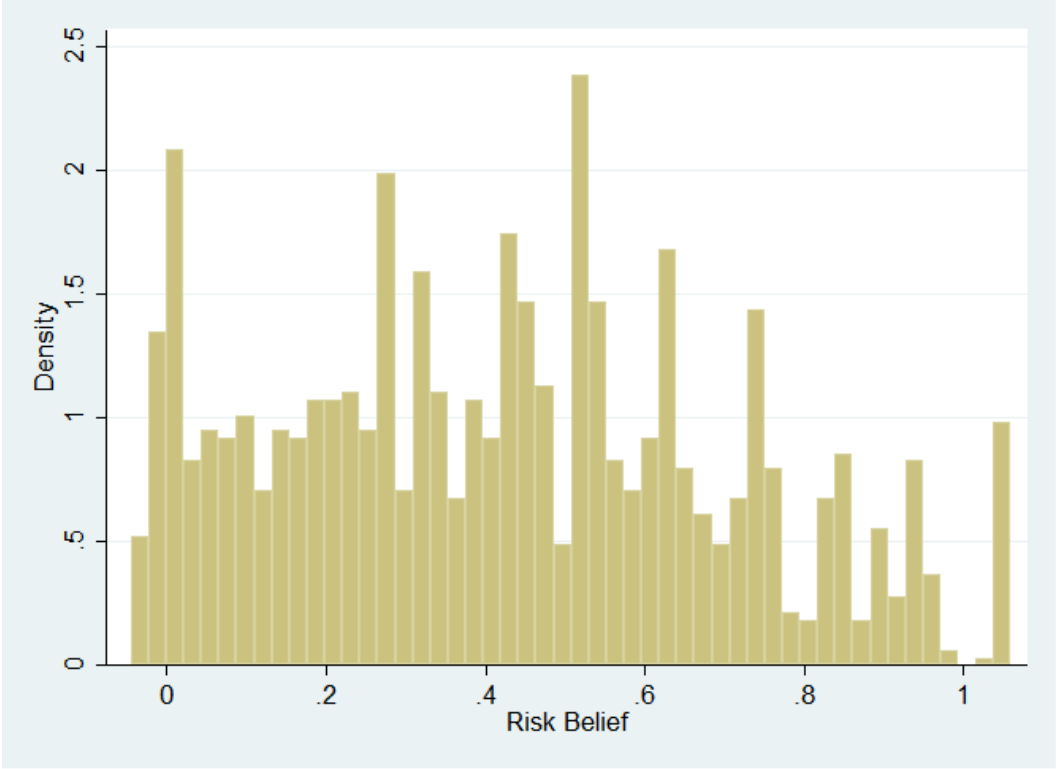
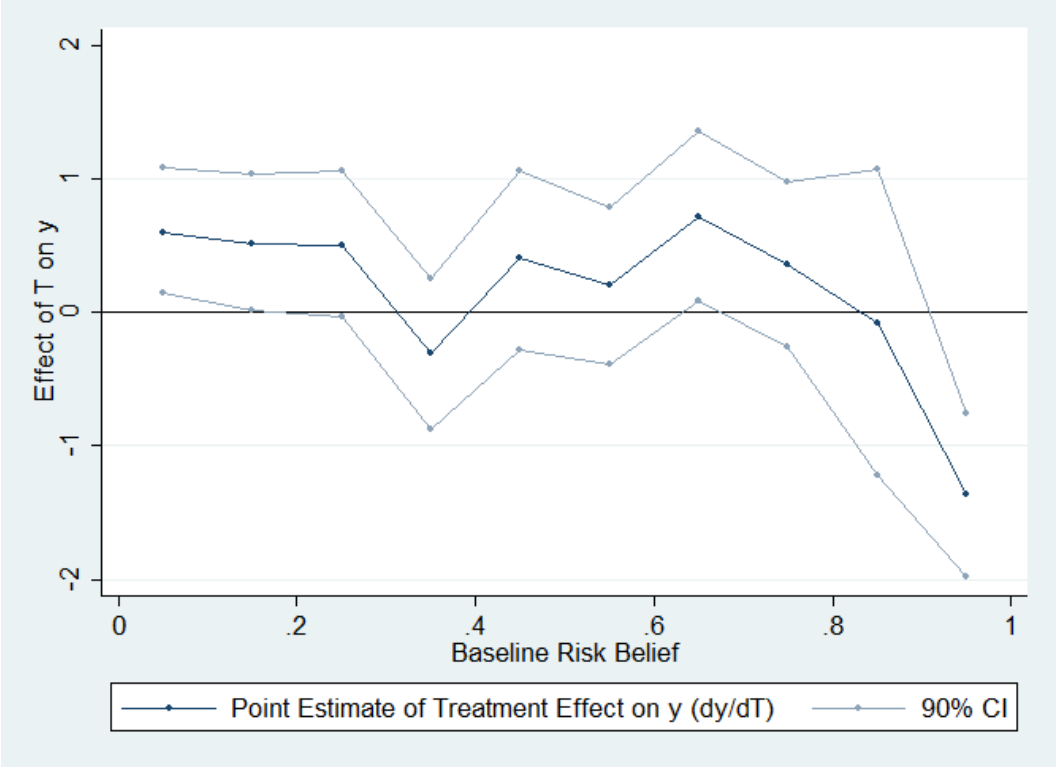


Figure 9: Effect of Treatment on Sex Acts in Past Week by Baseline Risk Belief

Panel A: Saturated Parametric Regression Approach



Panel B: Semiparametric Approach

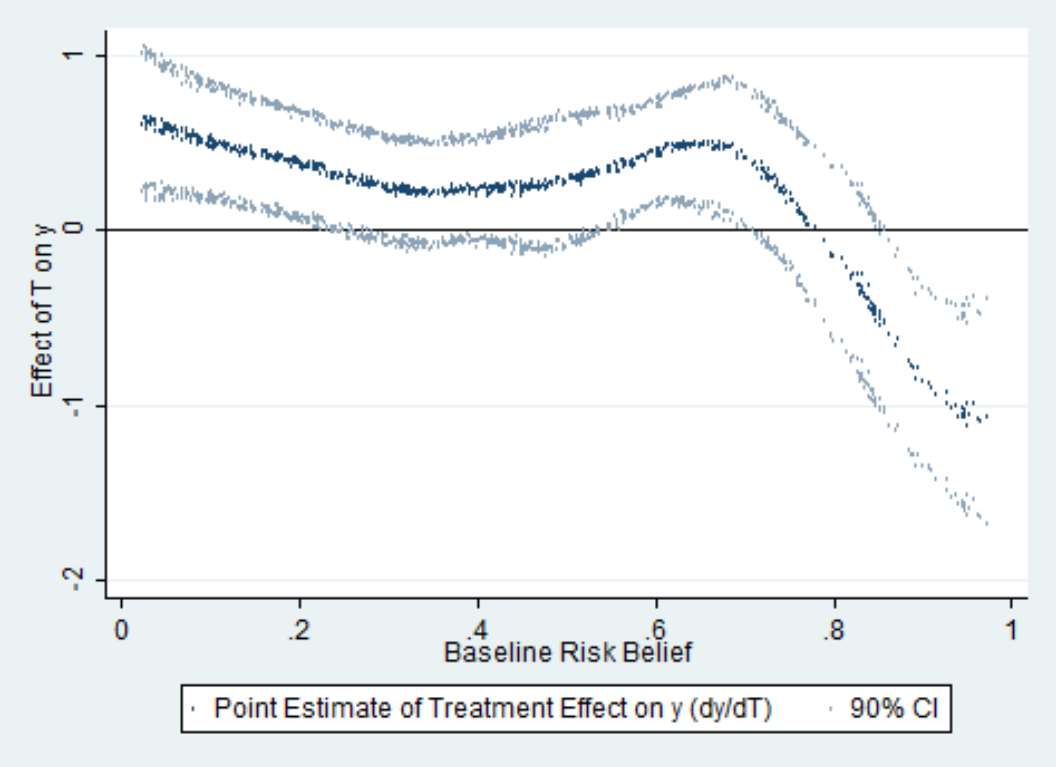
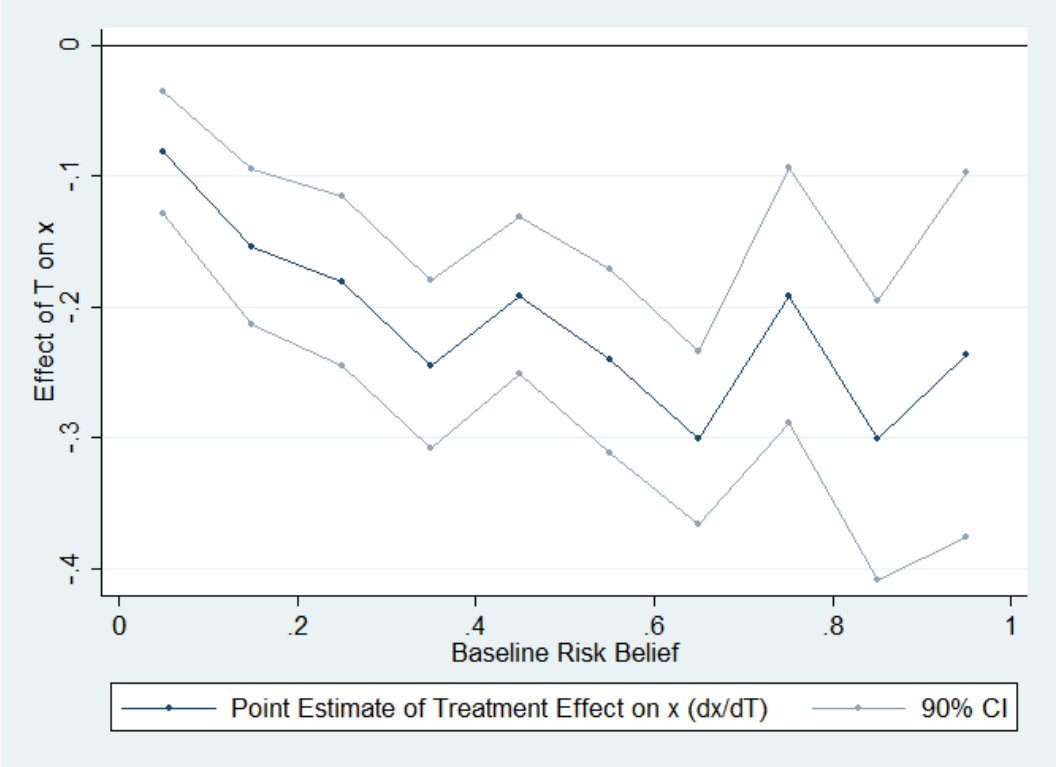


Figure 10: Effect of Treatment on Endline Risk Beliefs by Baseline Risk Belief

Panel A: Saturated Parametric Regression Approach



Panel B: Semiparametric Approach

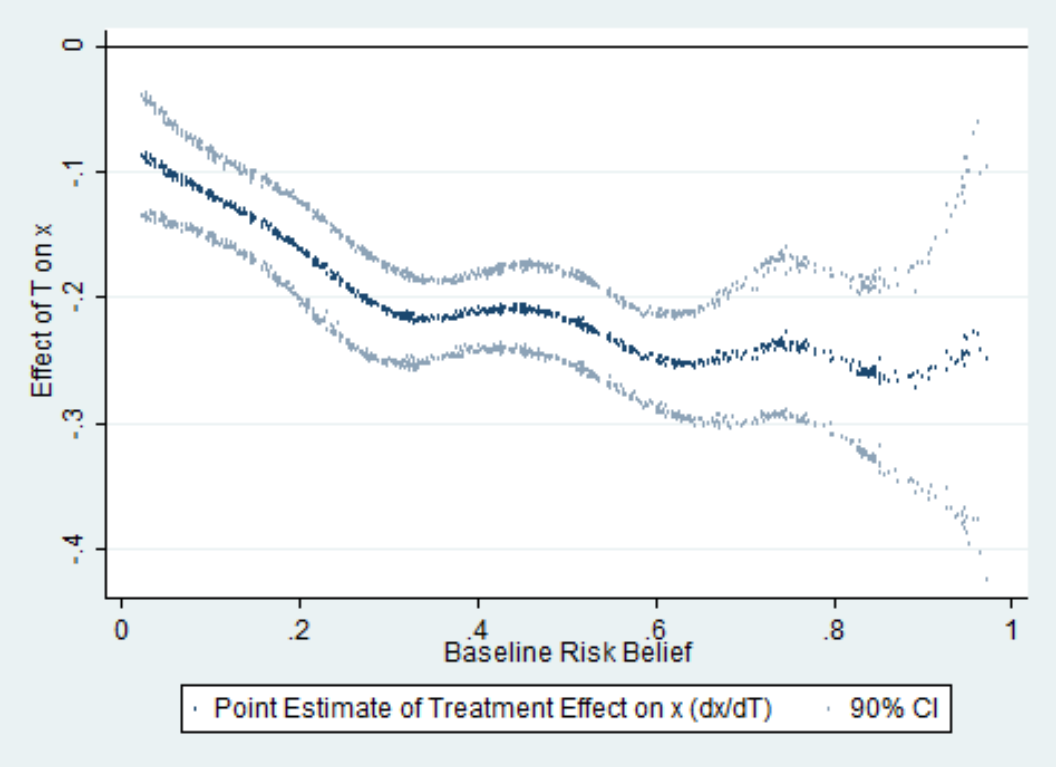
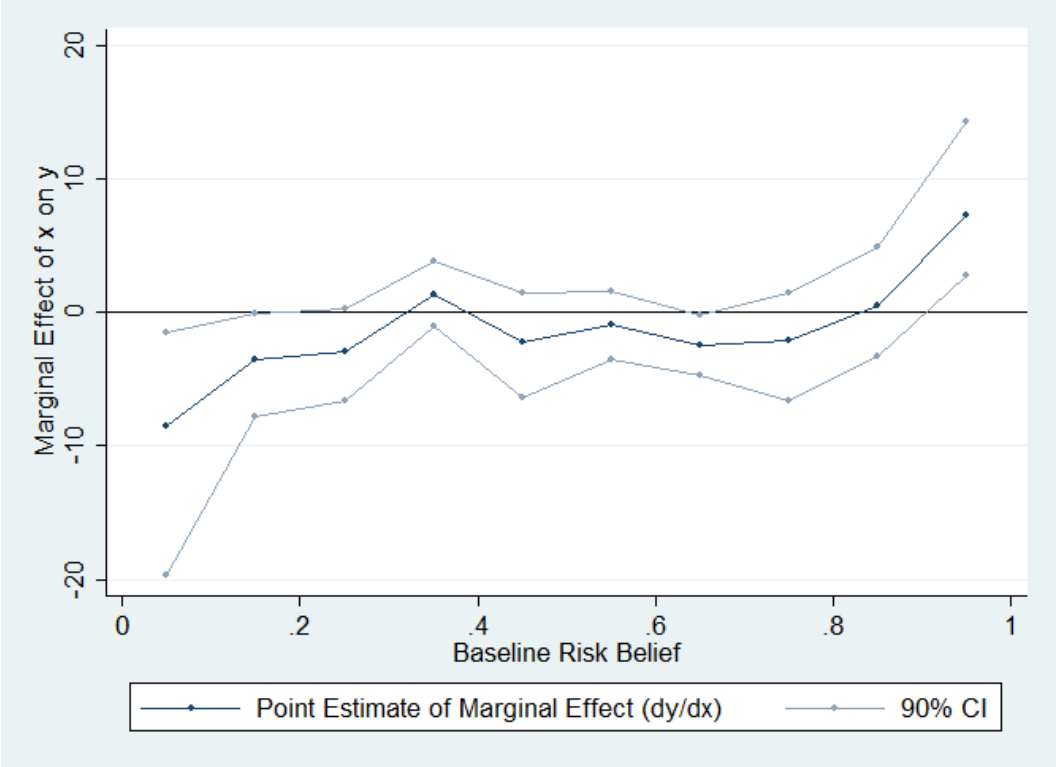


Figure 11: Effect of Endline Risk Beliefs on Sex Acts in Past Week by Baseline Risk Belief

Panel A: Saturated Parametric Regression Approach



Panel B: Semiparametric Approach

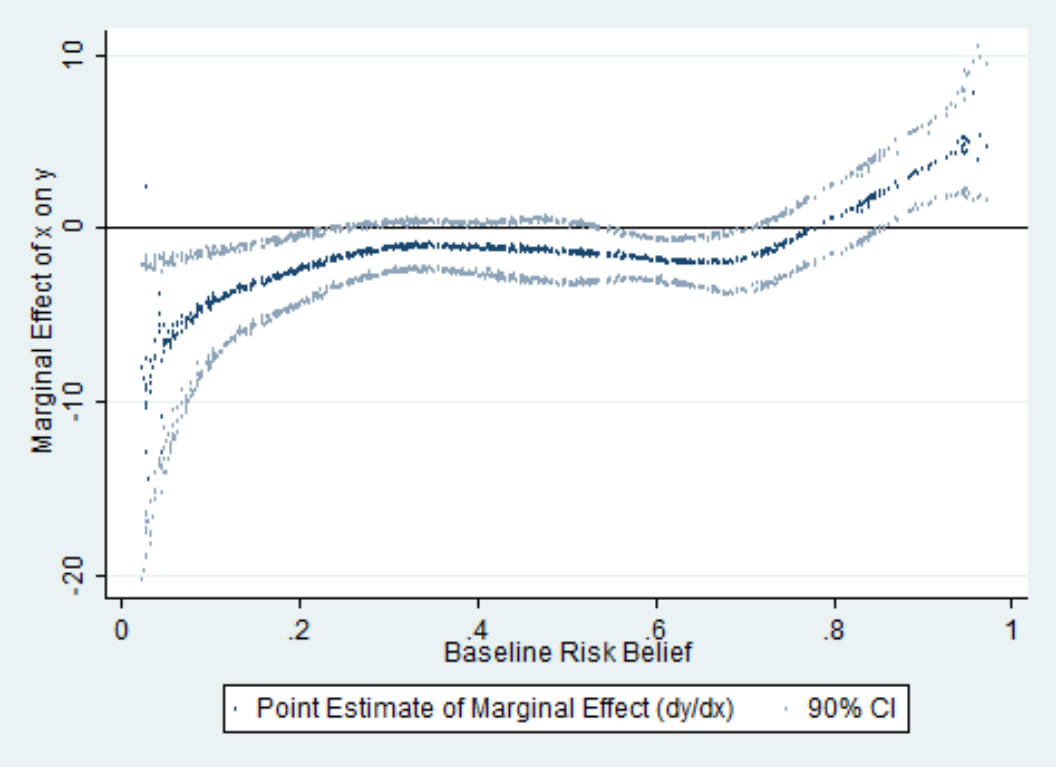


Figure 12: Effect of Treatment on Whether You Believe You are HIV-Positive Now, by Baseline Risk Belief

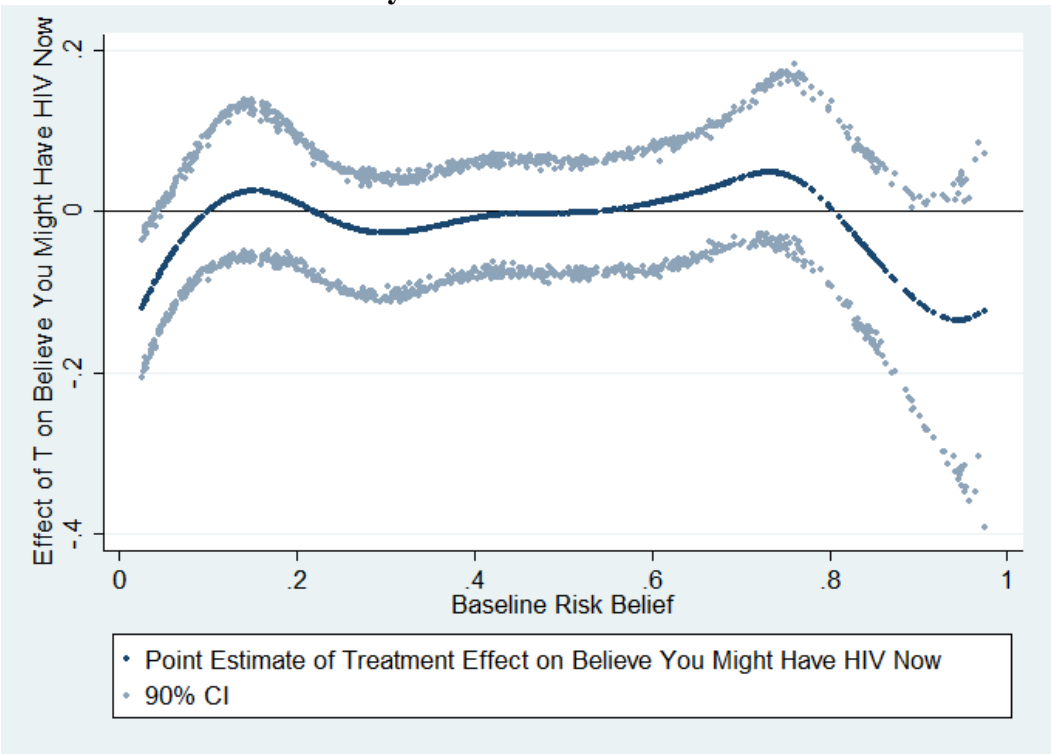


Figure 13: Effect of Treatment on Whether You Believe You Might Become HIV-Positive in the Future, by Baseline Risk Belief

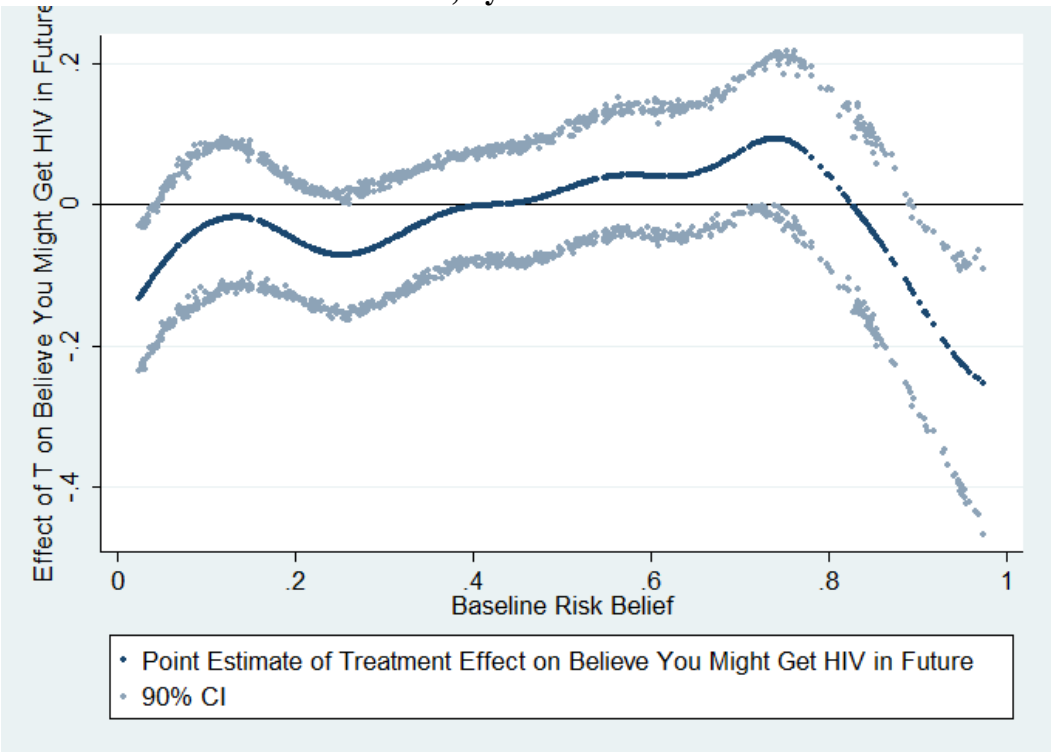


Table 1: Overall Sample Size

	Overall	Control	Treatment
Villages	70	35	35
0-2 km from a trading center	24	12	12
2-5 km from a trading center	24	12	12
5+ km from a trading center	22	11	11
Respondents	1292	645	647
Men	556	274	282
0-2 km from a trading center	191	98	93
2-5 km from a trading center	192	90	102
5+ km from a trading center	173	86	87
Women	736	371	365
0-2 km from a trading center	251	128	123
2-5 km from a trading center	238	118	120
5+ km from a trading center	247	125	122

Table 2: Attrition Analysis

	(1)	(2)
Treatment	0.02	0.02
	(0.02)	(0.02)
Constant	0.85***	0.81***
	(0.02)	(0.13)
Other controls		X
Observations	1503	1484

Notes: Sample is 1503 sexually-active adults who were successfully interviewed at baseline. Heteroskedasticity-robust standard errors, clustered by village, in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Other controls include (baseline-observed) marital status, age, age^2, whether respondent grew up in current village, education, total household size, number of living and desired future children, total number of media sources respondent uses at least once per month, respondent attractiveness rating (normalized by enumerator), total common assets owned by household, logged spending in the past month, logged income in the past month, and categorical indicators for sampling stratum, ethnic group, religion, and the week of the followup survey.

Table 3: Demographic Balance

	N	Overall	Control	Treatment	C-T
	(1)	(2)	(3)	(4)	(5)
Demographics					
Male	1292	0.43	0.42	0.44	-0.01
Married	1290	0.82	0.83	0.80	0.03
Age	1292	29.36	29.13	29.59	-0.46
Grew up in village where currently residing	1289	0.62	0.65	0.60	0.05
Years of education	1292	5.81	5.76	5.86	-0.10
Number of people in household	1292	4.95	5.04	4.87	0.17
Total children still living	1292	2.99	2.94	3.05	-0.11
Desired future children	1289	1.36	1.31	1.41	-0.09
# media sources† used at least monthly	1292	1.18	1.16	1.20	-0.04
# common assets owned by household	1291	4.40	4.54	4.26	0.28
Household cash income past 30 days (PPP USD)					
Baseline (non-overlapping periods)	1292	250.29	282.46	218.23	64.23**
Followup (observed at same time)	1292	190.28	201.94	178.66	23.29
Household expenditure past 30 days (PPP USD)	1292	292.70	292.39	293.01	-0.62
Religion					
Muslim	1292	0.07	0.09	0.06	0.02
Christian	1292	0.89	0.89	0.89	-0.01
Other	1292	0.04	0.03	0.05	-0.02
Ethnic Group					
Nyanja	1292	0.47	0.46	0.48	-0.02
Lomwe	1292	0.37	0.34	0.39	-0.05
Yao	1292	0.09	0.11	0.07	0.04
Chewa	1292	0.04	0.05	0.03	0.02
Other	1292	0.03	0.04	0.02	0.02

Notes: Sample is 1292 people from 70 villages for whom both baseline and followup surveys were successfully completed. Cluster-adjusted significance tests: * p< 0.1; ** p< 0.05; *** p<0.01. †: media sources include newspapers, radio, and television.

Table 4: Baseline Sexual Risk-Taking Measures

	N	Overall	Control	Treatment	C-T
	(1)	(2)	(3)	(4)	(5)
Panel A: Single-Question Recall					
Total lifetime sex partners	1288	3.34	3.12	3.56	-0.44**
Sex partners during past 30 days	1290	0.81	0.82	0.80	0.02
Any sex in the past 30 days	1281	0.73	0.74	0.73	0.01
Total sex acts during past 30 days	1281	7.37	7.48	7.27	0.21
Percent unprotected sex in the past 30 days	938	0.86	0.85	0.87	-0.01
7-Day Retrospective Sex Diary					
Respondent had any sex during past 7 days	1292	0.52	0.54	0.51	0.03
Total sex acts during past 7 days	1292	1.71	1.80	1.62	0.18
Percent unprotected sex during past 7 days	676	0.88	0.85	0.90	-0.05*
Respondent had sex with more than one partner in past 7 days	1292	0.01	0.03	0.01	0.02
Total sex acts with non-primary partners in past 7 days	1292	0.02	0.02	0.01	0.01
Percent unprotected sex with non-primary partners in past 7 days	15	0.47	0.40	0.60	-0.20

Notes: Sample is 1292 people from 70 villages for whom both baseline and followup surveys were successfully completed. Cluster-adjusted significance tests: * p< 0.1; ** p< 0.05; *** p<0.01.

Table 5: Baseline HIV Risk Beliefs

	N	Control	Treatment	C-T
	(1)	(2)	(3)	(4)
Panel A: Treatment Group Beliefs Measured at Baseline, Unadjusted				
HIV Transmission				
One Act				
Unprotected	1289	0.83	0.74	-0.09***
If using a condom	1291	0.12	0.10	-0.03**
One Year				
Unprotected	1284	0.93	0.88	-0.05***
If using a condom	1284	0.24	0.16	-0.08***
HIV Prevalence				
All local people of opposite sex	1279	0.53	0.51	-0.01
People respondent finds attractive	1276	0.55	0.52	-0.03*
Panel B: Treatment Beliefs Measured at Followup, Unadjusted				
HIV Transmission				
One Act				
Unprotected	1284	0.74	0.74	0.01
If using a condom	1286	0.08	0.10	-0.01
One Year				
Unprotected	1281	0.91	0.88	0.03**
If using a condom	1282	0.18	0.16	0.02
HIV Prevalence				
All local people of opposite sex	1270	0.48	0.51	-0.03
People respondent finds attractive	1269	0.46	0.52	-0.06***
Panel C: Treatment Beliefs Measured at Baseline, Adjusted for Non-Constant Linear Trends				
HIV Transmission				
One Act				
Unprotected	1289	0.79	0.79	-0.00
If using a condom	1291	0.14	0.15	-0.00
One Year				
Unprotected	1284	0.93	0.93	-0.00
If using a condom	1284	0.29	0.28	0.00
HIV Prevalence				
All local people of opposite sex	1279	0.54	0.53	0.01
People respondent finds attractive	1276	0.52	0.53	-0.00

Notes: Sample is 1292 people from 70 villages for whom both baseline and followup surveys were successfully completed. Cluster-adjusted significance tests: * p< 0.1; ** p< 0.05; *** p<0.01.

Table 6: Change in HIV Risk Beliefs due to Information Treatment

	N	Control	Treatment	C-T
	(1)	(2)	(3)	(4)
Risk Beliefs Measured at Followup				
HIV Transmission				
One Act				
Unprotected	1284	0.74	0.33	0.41***
If using a condom	1284	0.08	0.03	0.05***
One Year				
Unprotected	1284	0.91	0.52	0.38***
If using a condom	1284	0.18	0.09	0.09***
HIV Prevalence				
All local people of opposite sex	1269	0.48	0.32	0.16***
People respondent finds attractive	1269	0.46	0.41	0.06***

Notes: Sample is 1292 people from 70 villages for whom both baseline and followup surveys were successfully completed. Cluster-adjusted significance tests: * p< 0.1; ** p< 0.05; *** p<0.01.

Table 7: Impact of Information Treatment on Sexual Risk-Taking

	Any Sex in Past Week	# Sex Acts in Past Week	% Unprotected Sex Acts in Past Week	Sex Partners in Past 30 Days	Condoms acquired in past 30 days†	# condoms purchased	Overall Risk Index	Sex Diary Risk Index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment Group	0.05**	0.20*	0.01	-0.01	1.91***	0.1	0.02	0.04*
	(0.02)	(0.11)	(0.02)	(0.03)	(0.67)	(0.45)	(0.02)	(0.02)
Observations	1292	1292	493	1290	1283	1290	1292	1292
R-squared	0.2423	0.2376	0.1615	0.2471	0.1094	0.0484	0.2567	0.1226
Mean(Dep. Var)	0.39	1.26	0.88	0.80	3.99	5.15	-0.04	-0.05
SD(Dep. Var)	0.49	2.25	0.31	0.64	13.00	6.57	0.51	0.56
Effect Sizes Normalized by SD								
Point Estimate	0.10	0.09	0.03	-0.02	0.15	0.02	0.04	0.07
Top of 95% CI	0.20	0.18	0.19	0.06	0.25	0.15	0.14	0.16
Bottom of 95% CI	0.00	0.00	-0.13	-0.09	0.04	-0.12	-0.04	-0.02

Notes: Sample includes 1292 respondents who completed both baseline and followup surveys. Heteroskedasticity-robust standard errors, clustered by village, in parentheses. * p< 0.1; ** p< 0.05; *** p<0.01. All regressions estimated using ANCOVA with the exception of condom purchases (Column 6), where no data was collected at baseline, and all regressions include controls for sampling strata (distance category X gender). † Significant result for condoms acquired is not robust to excluding outliers, but is robust to Winsorizing.

Table 8: Marginal Effect of Risk Beliefs on Sexual Risk-Taking

	Any Sex in Past Week	# Sex Acts in Past Week	% Unprotected Sex Acts in Past Week	Sex Partners in Past 30 Days	Condoms acquired in past 30 days†	# condoms purchased	Overall Risk Index	Sex Diary Risk Index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: OLS Estimates								
Treatment Group	0.04	-0.01	-0.01	0.13**	-1.23	-0.8	0.04	0.08
	(0.04)	0.22	0.04	0.06	1.09	0.71	0.05	0.07
Observations	1252	1252	481	1250	1243	1250	1252	1252
R-squared	0.24	0.24	0.18	0.25	0.11	0.05	0.26	0.12
Panel B: 2SLS Estimates								
Endline Risk Belief	-0.26**	-1.17**	-0.03	0.05	-9.84***	-0.51	-0.11	-0.24*
	(0.12)	(0.55)	(0.09)	(0.14)	(3.71)	(2.25)	(0.11)	(0.12)
Observations	1252	1252	481	1250	1243	1250	1252	1252
R-squared	0.21	0.22	0.18	0.24	0.08	0.05	0.25	0.10
Panel C: 2SLS Estimates Normalized by Mean								
Point Estimate	-0.67	-0.93	-0.03	0.06	-2.47	-0.10	2.75	4.80
Top of 95% CI	-0.05	-0.06	0.18	0.40	-0.61	0.77	-2.75	-0.20
Bottom of 95% CI	-1.28	-1.79	-0.25	-0.28	-4.32	-0.97	8.25	9.80

Notes: Sample includes 1292 respondents who completed both baseline and followup surveys. Heteroskedasticity-robust standard errors, clustered by village, in parentheses. * p< 0.1; ** p< 0.05; *** p<0.01. All regressions include controls for baseline values of the outcome (with the exception of condom purchases (Column 6), where no data was collected at baseline), baseline values of risk beliefs, and sampling strata (distance category X gender). † Significant result for condoms acquired is not robust to excluding outliers, but is robust to Winsorizing.

Table 9: Heterogeneity in Impact of Information Treatment on Total Sex in Past Week

	(1)	(2)	(3)	(4)	(5)	(6)
Treatment (T)	0.201*	0.722***	0.273	0.108	0.291**	0.200
	(0.107)	(0.178)	(0.165)	(0.152)	(0.137)	(0.149)
T*(Baseline Risk Belief)		-1.187***				
		(0.365)				
T*(Male)			-0.167			
			(0.329)			
T*(Baseline Sex)				0.054		
				(0.074)		
T*(Ever Exposed to HIV)					-0.391	
					(0.312)	
T*(Some Chance I am HIV-positive)						-0.016
						(0.301)
Observations	1292	1275	1292	1292	1275	1277
R-squared	0.238	0.245	0.238	0.238	0.238	0.239

Notes: Sample includes 1292 respondents who completed both baseline and followup surveys. Heteroskedasticity-robust standard errors, clustered by village, in parentheses. * p< 0.1; ** p< 0.05; *** p<0.01. All regressions include controls for baseline values of the outcome, and sampling strata (distance category X gender). In each specification, the factor being interacted with the treatment dummy also enters into the regression in levels.

Table 10: Lifetime sex Partners by Effect of Information

	N	% of Total	Mean	SD
Negative Treatment Effect	90	14.17%	4.24	6.40
Positive Treatment Effect	545	85.83%	3.19	3.26