Navigating Ambiguity: Imprecise Probabilities and the Updating of Disease Risk Beliefs

Jason T. Kerwin and Divya Pandey*

March 20, 2025

Abstract

Probabilistic risk beliefs are key drivers of economic and health decisions, but people are not always certain about their beliefs. We study these "imprecise probabilities", also known as ambiguous beliefs. Imprecision is measurable separately from the level of risk beliefs, and higher imprecision leads to more updating of beliefs in response to a randomized information treatment. New information also causes changes in imprecision levels. We can map our data onto both a standard Bayesian model and a version that is designed to handle imprecise probabilities; these models match some features of our data but not all of them. Imprecise probabilities have important implications for our understanding of decisionmaking and for the design of programs intended to change people's minds.

Keywords: Development Economics, Health Economics, Imprecise Probabilities, Imprecision, Ambiguity, Ambiguous Beliefs, Epistemic Uncertainty, Knightian Uncertainty, Uncertainty, Risk Beliefs, HIV/AIDS
JEL Classification: O12, I15, D83

---

*Kerwin: Department of Applied Economics, University of Minnesota, 1994 Buford Avenue, St. Paul, MN 55108, USA, IZA, and J-PAL (jkerwin@umn.edu); Pandey: Department of Applied Economics, University of Minnesota, 1994 Buford Avenue, St. Paul, MN 55108. We thank Arnab Basu, Anja Benshaul-Tolonen, Daniel Chen, JD Denton-Schneider, Josh Dean, Adeline Delavande, Audrey Dorélien, Pamela Giustinelli, Paul Glewwe, Ben Hanowell, Cung Truong Hoang, Jessica Hoel, Kondwani Katundu, Hans-Peter Kohler, Caroline Krafft, Deborah Levison, Charles Manski, Maxwell Mkondiwa, Muhammad Meki, Ann Swidler, and seminar audiences at CSAE, the Economic Demography Workshop, the MLSFH Workshop, NEUDC, PAA, the Subjective Expectations Workshop, the University of Hawai'i, and the University of Illinois for helpful comments and suggestions. This research was supported in part by an NIA training grant to the Population Studies Center at the University of Michigan (T32 AG000221). The authors gratefully acknowledge support from the Minnesota Population Center (P2C HD041023) funded through a grant from the Eunice Kennedy Shriver National Institute for Child Health and Human Development (NICHD).

What is the chance that you will die within the next decade? Suppose you think the chance is 1 in 10. How *certain* are you about this chance? More broadly, can uncertainty about risk beliefs even be separated from the level of those beliefs? Conventional economic surveys treat these as indistinguishable: your risk belief perfectly captures your uncertainty. For example, if you think the risk is 10%, the standard deviation of this Bernoulli random variable is $\sqrt{0.1 * (1 - 0.1)} = 0.3$. That is, if you think the chance is 0% then you are certain you will not die in the next 10 years, and if you think the chance is 100% then you are certain that you will; likewise, uncertainty is maximized at a subjective probability of 50%. For binary variables, uncertainty is thus treated as a mechanical function of the level of risk beliefs. For continuous variables such as earnings, however, beliefs about levels have no such mechanical connection with uncertainty. Surveys often measure these beliefs by asking people not just about what they think their earnings will be next year, but also about the chance that their earnings will fall into different ranges (Dominitz 1998, Manski 2004).

In this paper, we study second-order uncertainty in probabilistic risk beliefs such as the chance of death or job loss. That is, we explore how uncertain people are about their probabilistic beliefs, which is often referred to as "ambiguity". Because "ambiguity" has a variety of meanings in economics, we use the term "imprecision" to refer to this concept.[1] This same idea has previously been explored in separate lines of work in the economics of subjective beliefs (e.g. Giustinelli, Manski, and Molinari 2022*a*) and in cognitive uncertainty (e.g. Enke and Graeber 2023 and Charles, Frydman, and Kilic 2022). Imprecision has been measured across many disciplines with varied approaches, including probabilistic risk analysis in engineering (Paté-Cornell 1996), confidence-based assessments in psychology (Yaniv and Foster 1995), and scenario-based modeling in climate change research (Cameron 2005).

We show that imprecision in probabilistic risk beliefs is crucial for addressing a key policy question: how can we convince people to believe new information? We define and measure imprecision separately from the level of risk beliefs. To take a specific example, one person who thinks they have a 10% chance of dying within the next decade could hold that belief with certainty, while another person with the same

[1] Specifically, we use "imprecision" to refer how uncertain probability beliefs are, and "imprecise probabilities" to describe probability beliefs with any imprecision.

point estimate thinks that the chance might range from 1% to 20%. Our measure of imprecision is the width of this range that people place around their probability estimates. We show that the width of this range is predictive of how willing people are to update their beliefs in response to new information.

While the distinction between imprecision and the level of one's risk beliefs is rarely drawn in economic research, it can be modeled using modifications of Bayesian models of belief updating. We show that both a standard Bayesian model and a "Robust Bayes" variation (based on Walley 1991) designed to handle imprecise probabilities can be mapped onto data about imprecision in risk beliefs. For both models, we use Beta-distributed priors over the risk belief variable. The two approaches give somewhat different results. The standard Bayesian model predicts that the extent of risk belief updating is proportional to how imprecise people's beliefs are: people who are less certain about their beliefs update more. In contrast, in the Robust Bayes approach, we find the counter-intuitive result that the extent of updating is independent of baseline imprecision levels. Both models yield a counter-intuitive prediction for changes in posterior imprecision—implying that they should depend only on the amount of new information provided, and not on the difference between the prior and the risk level in the new information.

To test these theoretical predictions we use data from a randomized experiment about HIV transmission risks in southern Malawi (Kerwin 2025). Perceived HIV transmission risks are extremely high among Malawians: the average respondent in our sample believed that the transmission rate from regular unprotected sex with an infected partner is 90% per year even though the true transmission rate is just 10% (Wawer et al. 2005, Malawi National AIDS Commission 2009). The experiment employed an information treatment that provided half of the participants with information about the actual HIV transmission risks at the end of the baseline survey.

We measure imprecision in risk beliefs by allowing survey respondents to indicate it for themselves. For risk belief questions that elicit a probability, our survey also asks for a range: what is the lowest the probability could be, and what is the highest it could be? This approach follows Manski (2004)'s suggestion that economists elicit a range of probabilities to capture imprecision in probability beliefs. These questions about probability ranges were asked as follow-ups for every risk belief question, allowing us to characterize imprecise probabilities for our entire sample. We focus our analysis in this paper on beliefs about the annual transmission rate of HIV, because the information treatment used annual risk beliefs. The predicted patterns of updating are most clear for this measure, since both the information treatment and the risk belief variable are about the same risk.

Measured levels of imprecision in our sample are relatively low. Focusing on beliefs about annual HIV transmission risks, the majority of our sample (76.8%) report a range of zero, indicating no imprecision. If we include the respondents with zero imprecision in our analysis, the average reported range has a width of just 5 percentage points (around an average risk belief of 90%). Among people who report a non-zero range, the average range is 21 percentage points wide. The low rates of imprecision in our sample are likely because HIV is a well-understood disease in Malawi—in sharp contrast to Delavande, Bono, and Holford (2021) whose document much higher rates of imprecision with regard to the risks of COVID-19, a novel health risk.

Our first key finding is that imprecision in probability beliefs can be measured separately from the levels of risk beliefs, and is not just a mechanical function of risk belief levels. Imprecision shows a marked inverse U-shaped relationship with the levels of risk beliefs—individuals holding the lowest and highest risk beliefs display the least imprecision. We eliminate the possibility of this inverse U-shaped relationship being the mechanical result of the ranges being bounded above at 100% and below at 0%. We measure the range above and below the point estimate of the belief, and find that it contracts on the non-bounded side as well. We also find that imprecision has limited correlations with individual characteristics: there are no strong correlations with age, education, or numeracy and also no significant differences between the magnitude of imprecision reported by male and female respondents.[2] We do find, however, that individuals who report "rounded probabilities" display significantly higher imprecision; their reported range of risk beliefs is 4 percentage points wider, an 80% increase relative to the mean.[3]

[2] The lack of correlation between imprecision and education is not due to a lack of understanding. There is a correlation between the level of risk beliefs and education, and previous work has shown that even respondents in developing-country settings with low formal education levels can understand subjective risk belief questions (Delavande, Giné, and McKenzie 2011).

[3] This finding is consistent with Giustinelli, Manski, and Molinari (2022*a*) and Giustinelli, Manski,

Second, we find that imprecision predicts a greater propensity to update one's beliefs in response to new information. The effects of the randomized information treatment on endline HIV transmission risk beliefs were significantly larger for people with higher baseline imprecision. A 10 percentage-point increase in imprecision increases the magnitude of the effect of the treatment by 2.5 percentage points. This is a fairly substantial difference: the average effect of the information treatment is about 35 percentage points. The finding that belief updating is proportional to imprecision is in line with the main prediction of our standard Bayesian model, and contradicts what the Robust Bayes model predicts.

Third, we show that new, credible information not only shifts the level of risk beliefs but also affects how imprecise people's beliefs are. The average effect of the information treatment on endline imprecision is close to zero, but this effect varies sharply by people's level of baseline risk beliefs: it is positive for people with high baseline risk beliefs and negative for those with low baseline beliefs. This is consistent with the fact that the information was most at odds with the priors of people with high baseline risk beliefs. Participants with the highest levels of HIV transmission risk beliefs faced the largest shock to their priors and thus increased their imprecision after the information treatment, while those with the lowest levels of risk beliefs did not show any change in their imprecision (because they faced the smallest shock to their priors). This finding is intuitive: people with low priors did not receive any new information and thus did not become any more or less uncertain about their beliefs, while people with high priors learned that their priors were wrong and thus became less sure about their beliefs. However, it conflicts with the predictions of both of our models, which suggests that a different modeling approach may be more appropriate here.

Our belief updating results withstand numerous robustness checks. Most crucially, they hold even if we control for treatment effect heterogeneity by the level of HIV risk beliefs, and allow that heterogeneity to be non-linear. This is important because even with no imprecision at all, Bernoulli random variables have an inverse U-shaped relationship between the expectation and the variance, which may also drive imprecision in beliefs. Our results are also robust to controlling for the interactions between

and Molinari (2022*b*), who document extensive rounding of probabilistic risk beliefs in the US Health and Retirement Study.

the treatment and our full list of baseline covariates. Moreover, they are robust to using alternative definitions of the imprecision variable. However, our findings hold only for annual (and not per-act[4]) HIV transmission risks. The most likely reason for this is that the information treatment was about annual HIV risks. While we would expect this treatment to lead to updating about per-act risks (and it does), the role of imprecision in this updating process should be attenuated.

While relatively little empirical work in economics distinguishes between the levels of risk beliefs and imprecision in those beliefs, there is a long theoretical tradition of drawing this distinction starting with Knight (1921). Researchers have used many names for the concept, ranging from "Knightian uncertainty" (Nishimura and Ozaki 2004), "ambiguity" (Sinz et al. 2008), and "epistemic uncertainty" (Fischhoff and Bruine De Bruin 1999), to simply "unreliable" and "uncertain" probabilities (Gärdenfors, Peter and Sahlin, Nils-Eric 1982, Schoemaker 1991). These differing terms can impede clarity about what is being discussed: "ambiguity" has other uses in economics, and "Knightian uncertainty" implies that the uncertainty cannot be quantified at all.[5] We follow Giustinelli, Manski, and Molinari (2022*a*)'s terminology of "imprecise probabilities" to refer to uncertainty in probabilistic risk beliefs, which often arises due to a lack of knowledge.[6]

Our paper contributes to a nascent literature in economics on measuring imprecision in dichotomous risk beliefs. Imprecision in subjective beliefs (as well as cognitive uncertainty in decision-making) has more commonly been elicited in the context of continuous variables such as earnings (e.g. Dominitz and Manski 1994). Recent research has begun to collect measures of uncertainty in dichotomous risk belief variables as well: Giustinelli, Manski, and Molinari (2022*a*) do so for dementia risk beliefs, while Hoel et al. (2022) do so for beliefs about fertilizer quality. Finally,

[4] The per-act HIV transmission risk is the probability of acquiring HIV from an infected individual from a single sex act (CDC 2019).

[5] Camerer and Weber (1992) discuss various ways of defining the concept of ambiguity. One common use of the term in economics is as a synonym for Knightian uncertainty (see e.g. Eichberger and Kelsey 2002). A related concept is "ambiguity aversion", which is a distaste for Knightian uncertainty (Machina and Siniscalchi 2014; Aydogan, Berger, and Bosetti 2023). Hoel et al. (2022) use "ambiguous beliefs" to refer to imprecise probabilities. "Strategic ambiguity" is used in game theory and international relations to refer to the intentional creation of uncertainty to achieve strategic aims (Baliga and Sjöström 2008).

[6] This term comes from an extensive literature on the concept of imprecise probabilities from the field of statistics (Walley 1991).

in a closely related paper on the elicitation of imprecise probabilities, Delavande, Bono, and Holford (2021) use an approach similar to ours to measure imprecision (referred to as "perceived ambiguity" in their paper) in the risk beliefs of COVID-19 infection and transmission, held by a sample of university students in the UK. We build on this previous work by collecting imprecision for all members of a representative sample, using a method that can easily be added to existing questions about probabilities (simply asking for a range). We also show that imprecision correlates with the level of risk beliefs, as well as with specific answers of 50%. However, we show that our imprecision measure can be distinguished from the level of beliefs as well as respondents' demographic characteristics. A separate line of work on "cognitive uncertainty" in decision-making measures the subjective probability that the true probability lies within a certain range (Enke and Graeber 2023). For instance, Enke and Graeber capture "cognitive uncertainty" by asking participants in belief updating experiments how certain they are that their posterior belief is within a two percentage-point window around their stated belief.[7] In contrast, we elicit the maximum width of a window around people's beliefs. More importantly, unlike Enke and Graeber we are able to rule out the possibility that the U-shaped relationship between risk beliefs and imprecision is driven by mechanical effects.

These results also contribute to the limited literature on belief formation and updating under ambiguity (Bachmann et al. 2020) by showing that imprecision plays an important role in the updating of risk beliefs. Related work on updating risk beliefs (Delavande 2008) studies how women update their beliefs about the effectiveness of contraceptives, showing that the strength of beliefs plays a role in the updating process. Our paper shows that a simple metric of imprecision in risk beliefs—the range that respondents put on a probability—matters for how much they update their beliefs in response to new information.[8] These findings also help demonstrate that our measure of imprecision captures a real aspect of people's reasoning and decision-making about risks. Overall, these results shed light on how to convince

[7] Other research has also studied cognitive uncertainty about continuous variables (e.g. Yang 2023 and Charles, Frydman, and Kilic 2022), rather than about the sorts of probabilistic risks that we examine in this paper.

[8] While we focus on imprecision in priors and are unable to distinguish imprecision from the noise in the processing of information or gauge the strength of the information in updating people's beliefs (see Augenblick, Lazarus, and Thaler 2023).

people to change their minds, which is of critical importance for issues ranging from disease control to economic policy.

This paper contributes to the broader literature on imprecision in people's beliefs and its role in belief updating in three ways. First, we integrate the literature on the elicitation of imprecision in subjective risk beliefs with a separate body of work on the role of cognitive uncertainty in belief updating. Giustinelli, Manski, and Molinari (2022*a*) show that it is possible to elicit imprecise probabilities for risk beliefs, while Enke and Graeber (2023) find that cognitive uncertainty plays a role in belief updating. Our own study brings these two concepts together by showing that imprecision in risk beliefs plays a role in belief updating. Second, our sample and context distinguish our study from the existing literature in two important ways: we study imprecision in a real-world setting, using a representative sample of adults from a low-income country. For example, Delavande, Bono, and Holford (2021) study college students in the UK, Giustinelli, Manski, and Molinari (2022*a*) use a sample of older Americans, and Bachmann et al. (2020) study the managers of firms in Germany. There are also studies that examine ambiguity in lab settings or online surveys, also in rich countries (e.g., Enke and Graeber 2023). By studying imprecision in HIV transmission beliefs in Malawi our paper provides the first evidence on imprecision in health risk beliefs in a low-income setting, building on previous research that studies the level of risk beliefs in the developing world (Delavande 2014, Delavande 2022). Using a simple measure to elicit imprecision we show that it is possible to elicit imprecision and study its role in updating in a context with low education levels and limited resources. Moreover, unlike most previous studies, we work with a probability sample of the entire prime-age adult population of the area in question. The study that comes closest to ours in this respect is Giustinelli, Manski, and Molinari (2022*a*), which uses the Health and Retirement Study (HRS) sample of older Americans. Third, our study contributes to the limited literature on the Bayesian updating of risk beliefs in the context of developing countries, by examining how actual updating processes compare to a Bayesian benchmark.

1 A Model of Risk Belief Updating

To understand how imprecision might affect the updating of risk beliefs, we develop two theoretical models—both related to the traditional Bayesian updating model. Standard Bayesian models have been used to model updating of risk beliefs (e.g. Viscusi and O'Connor 1984, Hakes and Viscusi 1997, Cameron 2005, Rheinberger and Hammitt 2018) and find that people broadly update their risk beliefs according to Bayes' rule.[9] In the specific context of developing countries, Bayesian updating has been used to model updating of beliefs primarily in the agricultural sector: Lybbert et al. (2007) model pastoralists' beliefs about future rainfall in Ethiopia and Kenya, while Hoel et al. (2022) model farmers' beliefs around fertilizer quality in Tanzania and Uganda, and Maertens, Michelson, and Nourani (2021) model farmers' beliefs around agricultural technologies in Malawi.

Building on this literature, we expect that the study participants in Malawi update their HIV risk beliefs similarly to people in other contexts. In our first model in Section 1.1, we map a standard Bayesian model with a Beta-distributed prior onto imprecise beliefs. We define the level of risk beliefs and the imprecision in beliefs respectively as the Beta distribution's mean and dispersion parameters, and show how baseline risk beliefs and imprecision affect the updating of beliefs. Second, in Section 1.2, we use a "Robust Bayes" model; unlike the standard Bayesian model, this model is designed to represent imprecise probabilities and thus offers a tighter mapping with the imprecision measure from our survey data. It does this by modeling how the upper and lower values of the range of an agent's risk beliefs change in response to new information. In this model, we map the level of risk beliefs onto the midpoint of the range, and imprecision onto the width of the range. Note that our model focuses on the updating of risk beliefs, unlike Woodford (2020), which relates to the imprecision in how people perceive signals ("imprecise internal representations"). The theoretical literature on ambiguous and imprecise beliefs has generally focused on decision-making conditional on beliefs or on jointly modeling decisions and belief updating, rather than belief updating alone (see Ilut and Schneider 2023).

These two theoretical models give somewhat different results. Both versions of

[9] A related literature on *Bayesian persuasion* is based on the assumption that people do update their beliefs using Bayes' rule (see Kamenica 2019 for a review).

the model find that the more new information people receive, the more they update their risk beliefs. However, the standard Bayesian model predicts more updating in response to new information when people have higher levels of imprecision, i.e. people more uncertain about their risk beliefs update their beliefs more. This is consistent with the intuition that people with more imprecision are more "persuadable". The Robust Bayes model, on the other hand, does not yield this result. Finally, both models predict that the magnitude of the gap between the prior and the new information has no effect on posterior imprecision. This is somewhat counterintuitive: when new information conflicts with our priors, it should make us less certain about the veracity of our new belief.

The difference in the results of the two models comes from the way in which we treat the parameter that governs the dispersion of the prior, ν. In our standard Bayesian model, this "learning parameter" is taken to be equivalent to the range of people's prior risk beliefs. In the Robust Bayes approach, the range is handled separately from the learning parameter. Our empirical evidence suggests that the prior range does indeed reflect people's learning parameters.

1.1 Mapping a Standard Bayesian Model onto Imprecise Beliefs

This section shows that imprecise probabilities can be represented in a natural way by a standard Bayesian model of belief updating. Let $p \in [0, 1]$ represent a potential value of the risk belief variable. We can represent the agent's priors for p using a Beta distribution, which is the conjugate prior for a Bernoulli-distributed random variable. This choice means that the posterior distribution follows the same parametric form as the prior distribution, which greatly simplifies both computation and intuition (see Gelman et al. (2013)). The Beta distribution gives the probability density of p based on two shape parameters, $\alpha > 0$ and $\beta > 0$:

$$f(p; \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}$$

where $\Gamma(\cdot)$ is the Gamma function. The shape parameters, α and β, can take on any positive real value, but if they are natural numbers then they correspond to the number of previously observed successes and failures.

The mean and variance of p are

$$\mathbb{E}[p] = \frac{\alpha}{\alpha+\beta}$$
$$\text{Var}[p] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

We follow Ferrari and Cribari-Neto (2004) in re-parameterizing the distribution in terms of its mean and dispersion. In particular, we define:

$$\mu = \frac{\alpha}{\alpha+\beta}$$
$$\nu = \frac{1}{1+\alpha+\beta}$$

Then $\mathbb{E}[p] = \mu$, while $\text{Var}[p] = \nu(\mu(1-\mu))$. The dispersion parameter, ν, controls the spread of the distribution: the larger ν is, the higher the variance of p. Thus μ corresponds to the level of risk beliefs in our data, and ν corresponds to the imprecision about those beliefs. A key feature of our empirical setting is that we observe both a point estimate and a range for all the respondents in our sample (although the ranges have zero width for much of our sample). The fact that we have both point estimates and ranges allows us to map the point estimate onto μ and the range onto ν.

Suppose the agent observes new information about the risk with n observations (i.e. the amount of new data available to the agent) and a risk x. For example, if the risk is the chance of a coin landing heads up, n is the number of coin flips the agent observes and x is the probability that each flip comes up heads. In our empirical setting, n is the number of initially HIV-negative people in the study that the information treatment described, which is 100, and x is the annual rate of HIV transmission, which is 10%.

Given this new information, the posterior distribution of p is given by $f'(p; \alpha, \beta | x, n) = f(p; \alpha + xn, \beta + n - xn)$. The reparameterized posterior shape parameters are

$$\mu' = (\alpha + xn)/(\alpha + \beta + n)$$
$$\nu' = 1/(\alpha + \beta + n + 1)$$

Our main interest is in how much the posterior mean of risk beliefs μ' updates in response to the risk x:

$$\begin{aligned}\frac{\partial\mu'}{\partial x} &= \frac{n}{\alpha+\beta+n} \\ &= \frac{n}{1/\nu+n-1} = \frac{n\nu}{1+n\nu-\nu}\end{aligned}$$

This expression is always positive. It says that the posterior expectation of p, μ', is an increasing function of the risk in the new information, x. We would thus expect larger treatment effects on μ' for people with bigger gaps between the prior μ and the new information x.

Next, we want to know how changes in the prior dispersion parameter ν affect the extent of updating of the posterior mean:

$$\frac{\partial}{\partial\nu}[\frac{\partial\mu'}{\partial x}] = \frac{n}{(\nu(n-1)+1)^2}$$

This expression is always positive, so the extent of updating is greater if ν is larger. This makes sense: larger ν means the variance of the distribution is larger, so people are less certain about their prior. This matches the main results in our paper. Figure 1 shows this result graphically: in Panel A, with tighter (more precise) priors, there is relatively little updating of beliefs, while the wider (more imprecise) priors in Panel B lead to more updating.

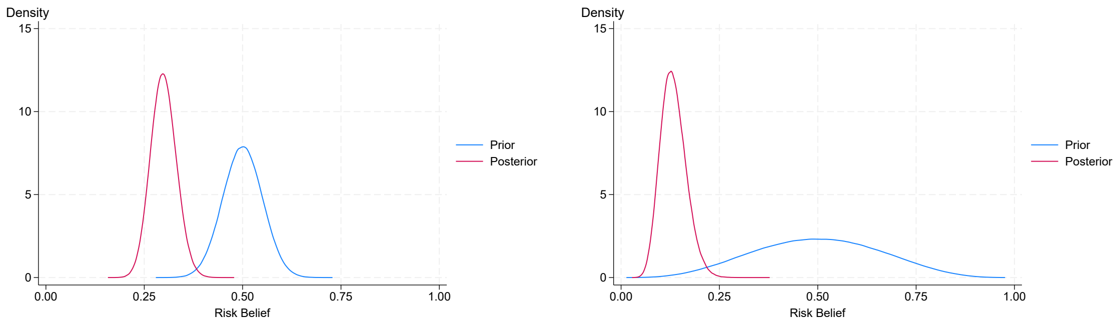
What about the effect of the amount of new information n on updating? This is

$$\frac{\partial\mu'}{\partial n} = \frac{\alpha(x-1)+\beta x}{(\alpha+\beta+n)^2}$$

The sign of this expression is ambiguous: it is positive if $\alpha x+\beta x > \alpha$, so if $x > \alpha/(\alpha+\beta)$, i.e. if $x > \mu$. So more information increases the posterior mean if the risk in the new information is above the prior mean, and decreases it if the risk is below the prior mean. This is exactly what we would expect intuitively: more n implies more updating. In our empirical application, the risk in the new information is below the prior mean for most respondents.

We can also derive predictions for how new information affects the posterior dis-

Figure 1
Updating of Risk Beliefs in Response to New Information
by Baseline Imprecision



Panel A: Low Imprecision, $\nu = 0.01$ **Panel B:** High Imprecision, $\nu = 0.1$

Notes: Graphs show the prior and posterior distributions of risk beliefs using a Beta distribution with expected value $\mu = 0.5$ and two different levels of the dispersion parameter ν. The new information has $n = 100$ observations and an average risk of $x = 0.10$, i.e. 90 failures and 10 successes.

persion parameter:

$$\frac{\partial \nu'}{\partial x} = 0$$

$$\frac{\partial \nu'}{\partial n} = -\frac{1}{(\alpha + \beta + n + 1)^2} = -\frac{1}{(\frac{1}{\nu} + n)^2} < 0$$

$$\frac{\partial}{\partial \nu}[\frac{\partial \nu'}{\partial n}] = -\frac{2\nu}{(1 + n\nu)^3} < 0$$

$$\frac{\partial}{\partial \mu}[\frac{\partial \nu'}{\partial n}] = 0$$

Posterior imprecision does not depend on the level of the risk in the new information, x, only on how much information there is, n. Larger amounts of information lead to lower posterior imprecision. This effect is larger for people with prior imprecision but is not affected by the level of prior beliefs. The latter prediction is somewhat counterintuitive: we would expect imprecision to increase if there is a larger gap between the prior and the new information. Intuitively, we would expect imprecision to increase by more for people who experience larger shocks to their priors. This would imply that $\frac{\partial^2}{\partial \mu \partial c}[\frac{\partial \nu'}{\partial n}] > 0$, but the model predicts this cross-partial derivative will be zero.

To summarize, our standard Bayesian model makes two key predictions. First, the extent of posterior belief updating is higher for people with more prior imprecision. Second, changes in posterior imprecision depend only on the amount of new information and not on the difference between people's priors and the risk level in the new information.

1.2 "Robust Bayes" Model

Now we will approach this modeling problem in a slightly different way, using an approach to inference that is designed to handle imprecise probabilities (Walley, 1991, pp. 217–221). This specific approach is part of a set of methods often referred to as "Robust Bayes", although Walley does not use that term himself. He defines three key parameters of a Beta-Bernoulli model with imprecise beliefs; we will alter his notation slightly to better match our own from above. Instead of a single expected value for the Beta distribution, define $\overline{\mu}$ and $\underline{\mu}$ as the top and bottom of the agent's range of beliefs. As above, let ν be the dispersion parameter. However, following Walley, we will now interpret this as the *learning parameter*: it determines how much the agent updates their beliefs in response to new information. This contrasts with the approach in Section 1.1, where we interpreted ν as capturing the imprecision in the agent's beliefs.

Posterior beliefs after observing n trials with a success rate of x are given by

$$\overline{\mu}' = \frac{\nu\overline{\mu} + nx}{\nu + n}$$
$$\underline{\mu}' = \frac{\nu\underline{\mu} + nx}{\nu + n}$$
$$\nu' = \nu + n$$

So the agent updates their range of beliefs from $[\underline{\mu}, \overline{\mu}]$ to $[\underline{\mu}', \overline{\mu}']$. In other words, the agent updates their beliefs in the same way as a normal Bayesian with Beta priors, but there are two different expectations instead of just one.

To map this setup onto our experiment, we need to define two additional parameters. First, the setup above admits the boundaries of the range but not a point estimate, whereas all the respondents in our sample report point estimates as well. We define the point estimate of the probability as $\mu = (\underline{\mu} + \overline{\mu})/2$, i.e. halfway between the top and the bottom of the range.[10] Second, we define the width of the range as $\omega = \overline{\mu} - \underline{\mu}$.[11] We can use these two new parameters to reparameterize the model as follows:

$$\overline{\mu} = \mu + \omega/2$$
$$\underline{\mu} = \mu - \omega/2$$

And we can rewrite the posterior range boundaries as well:

$$\overline{\mu}' = \frac{\nu(\mu + \omega/2) + nx}{\nu + n} = \frac{\nu\mu + \nu\omega/2 + nx}{\nu + n}$$
$$\underline{\mu}' = \frac{\nu(\mu - \omega/2) + nx}{\nu + n} = \frac{\nu\mu - \nu\omega/2 + nx}{\nu + n}$$

We can use these updated values to define the posterior versions of μ and ω:

$$\begin{aligned}
\mu' &= (\overline{\mu}' + \underline{\mu}')/2 = (\frac{\nu\mu + \nu\omega/2 + nx}{\nu + n} + \frac{\nu\mu - \nu\omega/2 + nx}{\nu + n})/2 \\
&= \frac{\nu\mu + nx}{\nu + n} \\
\omega' &= \overline{\mu}' - \underline{\mu}' = \frac{\nu\mu + \nu\omega/2 + nx}{\nu + n} - \frac{\nu\mu - \nu\omega/2 + nx}{\nu + n} \\
&= \frac{\nu\omega}{\nu + n}
\end{aligned}$$

Now we can compute the key comparative statics. First, we want to know how the level of risk beliefs μ is affected by changes in the per-act risk x:

$$\frac{\partial \mu'}{\partial x} = \frac{n}{\nu + n}$$

How does the range of risk beliefs affect the updating of beliefs? Not at all:

$$\frac{\partial^2 \mu'}{\partial x \partial \omega} = 0$$

[10] This assumption is clearly false for many of our respondents, who report a point estimate of 100% but a non-zero range. We maintain this assumption for simplicity, but our results could be extended to allow for other weighted averages instead.

[11] If we were to treat ν as capturing the range in this model, then we would get identical results to those in Section 1.1. This is consistent with Walley's description of ν as a learning parameter: it governs the extent to which new information leads to updating of risk beliefs.

Similarly, we find the effect of the amount of new information n on the level of risk beliefs:

$$\frac{\partial \mu'}{\partial n} = \frac{(x-\mu)\nu}{(n+\nu)^2}$$

More information leads to higher posterior mean beliefs when the new data is above the prior mean, and lower posterior mean beliefs when the new data is below the prior. Once again, this comparative static is unaffected by the width of the range:

$$\frac{\partial^2 \mu'}{\partial n \partial \omega} = 0$$

What about the effects of the new information on the width of the range?

$$\begin{aligned}
\frac{\partial \omega'}{\partial x} &= 0 \\
\frac{\partial \omega'}{\partial n} &= -\frac{\nu\omega}{(\nu+n)^2} \\
\frac{\partial}{\partial \omega}[\frac{\partial \omega'}{\partial n}] &= -\frac{\nu}{(\nu+n)^2} \\
\frac{\partial}{\partial \mu}[\frac{\partial \omega'}{\partial n}] &= 0
\end{aligned}$$

So the width of the range is unaffected by the level of the risk in the new information, but is a declining function of the amount of new information. This latter effect is larger if the range is wider, but is unaffected by the level of the risk belief.

These results are somewhat counterintuitive. We would expect risk beliefs to update more in response to new information when the range is wider, since the width reflects the extent of people's uncertainty. This gap between intuition and the model differs from the way we interpreted the standard Bayesian model, wherein we found that higher imprecision led to more updating.

The lack of any effect of the level of the risk belief on the width of the posterior interval is also counterintuitive. What we would expect is that when new information conflicts with the prior, the width goes up (since people become less certain), and that when it is consistent with the prior the width goes down.

The relationship between the width of the posterior interval and the amount of new information is mostly consistent with intuition. More information makes people more sure, and this effect is larger if they are less sure *ex ante*. However, we would expect this pattern to depend on the level of the risk in the new information and it does not.

1.3 Updating when agents have zero imprecision

Both theoretical models give similar results for updating when we consider imprecision to be zero for agents. We first evaluate the comparative statics of interest under our Standard Bayesian model, if the dispersion parameter, ν, is zero. In this case the agent updates neither the posterior mean of risk beliefs, μ', nor the posterior dispersion parameter, ν', in response to the risk x:

$$\left.\frac{\partial \mu'}{\partial x}\right|_{\nu=0} = \frac{n0}{1 + n0 - 0} = 0$$
$$\left.\frac{\partial \nu'}{\partial x}\right|_{\nu=0} = 0$$

This makes intuitive sense, as people who are firm in their beliefs are less likely to update their original beliefs as well as the imprecision they hold around those beliefs.

Under our Robust Bayes model, the change in the posterior mean of risk-belief is independent of the width of the range, ω, so a limiting case gives the same result as shown in Section 1.2. The width of the range is also not affected by the level of risk in the new information:

$$\left.\frac{\partial \mu'}{\partial x}\right|_{\nu=0} = \frac{n}{\nu + n}$$
$$\left.\frac{\partial \omega'}{\partial x}\right|_{\nu=0} = 0$$

This runs counter to intuition, since we would expect people to update their posterior less when their prior is more certain.

2 Experiment and Data

Our data comes from a randomized experiment conducted in the Zomba district of Malawi from August to December 2012 (Kerwin 2025). The experiment, designed to study the effect of risk beliefs on risk-taking behavior, collected data on a representative sample of 1,503 sexually active individuals (stratified by gender) from 70 randomly selected villages (stratified by distance to the nearest trading center) from one sub-district of Zomba. The endline survey, conducted 1-4 months after the baseline, successfully followed up with 1,292 of these individuals. The resulting sample was balanced across study arms on observed baseline characteristics (Table 1). The average respondent in the study was 29 years old, married, and had completed 6 years of primary school.

Risk beliefs about HIV transmission in Malawi are extremely high. The average respondent in our study believed that the transmission rate is 90% per year, while the true rate is about 10% (Wawer et al. 2005, Malawi National AIDS Commission 2009). The randomized treatment in the experiment made use of this information from Malawi's National Aids Commission—a noted public health authority in the country—and provided information on the true rates of HIV transmission risk to half of the respondents at the end of the baseline survey. This information treatment took the form of a script and a set of visual aids. Treatment-group respondents were told about the Wawer et al. (2005) study of HIV transmission in Rakai, Uganda. Specifically, they were told that out of a sample of 100 couples in Uganda with one HIV-positive and one HIV-negative partner, 10 of the HIV-negative partners had contracted HIV after one year of regular unprotected sex (about three times per week).[12] Following Godlonton, Munthali, and Thornton (2016), the treatment-group baseline surveys were conducted after the completion of the control-group baseline surveys to minimize the risk of contamination of the control villages.

Subjective risk beliefs about HIV transmission risks were collected at both baseline and endline. The survey questions asked about the number of people out of a fixed denominator who would contract HIV under certain conditions; the questions were matched to the respondent's gender. For example, for unprotected annual risks, the

[12] For simplicity of explanation, the original numbers from Wawer et al. were rounded and the denominator was set to 100.

Table 1
Baseline Balance

	Ctrl. Mean (SD) (1)	Treat. Mean (SD) (2)	Diff. (p-val.) (3)	Obs (4)
Demographics				
Male	0.425	0.436	0.000	1,292
	(0.495)	(0.496)	(1.000)	
Married	0.829	0.803	-0.025	1,290
	(0.377)	(0.398)	(0.316)	
Age	29.133	29.589	0.465	1,292
	(8.417)	(8.333)	(0.339)	
Years of Education	5.758	5.858	0.097	1,292
	(3.347)	(3.484)	(0.723)	
Household size	5.039	4.870	-0.176	1,292
	(2.237)	(2.036)	(0.254)	
Spending in past 30 days	292.390	293.010	1.698	1,292
	(383.593)	572.544	(0.954)	
Assets owned	4.543	4.263	-0.277	1,291
	(2.644)	(2.537)	(0.185)	
Raven's score [0-3]	1.551	1.538	-0.019	1,291
	(0.989)	(1.002)	(0.766)	
Numeracy [0-3]	0.715	0.818	0.096*	1,292
	(0.929)	(1.007)	(0.095)	
Risk attitude	0.261	0.274	0.014	1,288
	(0.440)	(0.447)	(0.634)	
Christian	0.910	0.927	0.017	1,292
	(0.286)	(0.260)	(0.472)	
Muslim	0.085	0.060	-0.025	1,292
	(0.280)	(0.238)	(0.281)	
Sexual Activity				
Any Sex in Past Week	0.541	0.507	-0.036	1,292
	(0.499)	(0.500)	(0.111)	
Total Acts in Past Week	1.798	1.615	-0.185	1,292
	(2.471)	(2.380)	(0.155)	
Unprotected Acts in Past Week	1.569	1.471	-0.100	1,292
	(2.376)	(2.323)	(0.446)	
Sex Partners in Past 30 Days	0.818	0.797	-0.024	1,290
	(0.498)	(0.762)	(0.515)	
Condoms Acquired in Past 30 Days	4.739	3.530	-1.205	1,288
	(15.003)	(11.549)	(0.122)	
Years Sexually Active	13.100	13.204	0.117	1,275
	(8.279)	(8.603)	(0.815)	
Lifetime Sex Partners	3.117	3.557	0.414**	1,288
	(2.684)	(4.734)	(0.042)	
Any Chance of Having HIV	0.344	0.352	0.008	1,277
	(0.475)	(0.478)	(0.788)	
Overall Sexual Activity Index	0.028	-0.028	-0.059	1,277
	(0.997)	(1.003)	(0.266)	

Notes: Sample is 1,292 people from 70 villages who completed both baseline and endline surveys. Differences and p-values in column 3 are adjusted for stratification cell fixed effects and clustered by village: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

enumerators asked women "If 100 women, who do **not** have HIV, each have an HIV-positive sex partner for **one year**, and do **not** use condoms when having sex, how many of the women do you think will have HIV at the end of the year?" There were similarly structured questions for other kinds of risks. If the response to a survey question was 50, enumerators asked respondents whether they really thought the response was 50, or whether they were just not sure. In the latter case, respondents were asked for their best guess; we use these best guesses instead of the original responses.

Our measure of imprecision in probability beliefs comes from another follow-up question, asked immediately after each risk belief question. These questions asked respondents what was the lowest and highest that the probability of HIV transmission could possibly be. For example, if a respondent gave an initial response of 90% for the unprotected annual risk of HIV transmission, they might say the highest the answer could be is 95 and the lowest it could be is 75. The questionnaire design prevented logical inconsistencies in the responses: if the initial response was at the boundary of the parameter space (i.e. 0 or 100), the lower or higher question was skipped because there is no way for the answer to go any higher. In these cases we infer that the answer must be the same as the original risk belief response. See Figure A1 for an example of what the survey questionnaire looked like for unprotected annual transmission risks.

We construct an imprecision measure for each question as the range between what the respondent indicates is the highest and lowest possible value. This measure has a close mapping to the imprecision variable our Robust Bayes model from Section 1.2, which is the width of a range around the agent's mean risk belief. Almost 80% of our sample expresses no imprecision at all. One possible reason for this is that HIV has been common in Malawi for decades, so people's beliefs about the risks of contracting it may be ingrained. Across the entire sample (including the zeroes) the average respondent reports an imprecision of 5 percentage points for annual risk beliefs and 8 percentage points for per-act risk beliefs (Table 2). For people reporting a non-zero range, the average imprecision is 21 and 20 percentage points for annual and per-act risk beliefs, respectively (Table A1); this is very similar to the average ranges that Giustinelli, Manski, and Molinari (2022*a*) find for dementia risk beliefs. Non-responses were limited to 8 participants for both the annual risk beliefs as well as the corresponding imprecision measure. For per-act risk beliefs and the per-act imprecision, non-responses were recorded for 3 and 5 participants, respectively. While the survey collected data for both per-act and annual risk beliefs, our preferred specifications use annual risk beliefs since the information treatment itself provided annual transmission risks.

Table 2

Summary Statistics for Risk Beliefs and Imprecision

	N	Mean	SD	Percentiles										
				Min	1st	5th	10th	25th	50th	75th	90th	95th	99th	Max
HIV Transmission Rate Beliefs														
Annual	1,284	0.90	0.20	0.04	0.10	0.40	0.60	0.90	1.00	1.00	1.00	1.00	1.00	1.00
Imprecision	1,284	0.05	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.30	0.70	0.99
Per-act	1,289	0.78	0.30	0.00	0.03	0.10	0.22	0.60	0.98	1.00	1.00	1.00	1.00	1.00
Imprecision	1,287	0.08	0.16	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.30	0.40	0.80	1.00
HIV Prevalence Beliefs														
Attractive People	1,276	0.54	0.27	0.00	0.00	0.00	0.20	0.40	0.50	0.70	0.90	1.00	1.00	1.00
Imprecision	1,269	0.10	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.30	0.40	0.50	0.90
All People	1,279	0.52	0.30	0.00	0.00	0.05	0.10	0.20	0.50	0.80	0.90	0.97	1.00	1.00
Imprecision	1,272	0.09	0.13	0.00	0.00	0.00	0.00	0.00	0.05	0.14	0.27	0.38	0.60	0.89

Notes: Sample is 1,292 people from 70 villages who completed both baseline and endline surveys. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk.

3 Empirical Strategy

To test whether imprecise probabilities play a role in updating of risk beliefs, we estimate treatment effect heterogeneity by baseline imprecision. In our main regression (Equation 1), Y_{ij}^{el} indicates HIV risk beliefs for respondent i from village j at endline (el), and Y_{ij}^{bl} denotes risk beliefs measured at baseline (bl). The indicator variable T_j takes a value of one if village j was randomized into the information treatment and is zero otherwise. A_{ij}^{bl} is imprecision at baseline. The coefficient of interest, β_4, measures how much the treatment effect varies by baseline imprecision. Since A_{ij}^{bl} is correlated with Y_{ij}^{bl} we also control for heterogeneity by Y_{ij}^{bl} (β_5). We also control for stratification cell fixed effects $\boldsymbol{S_{ij}}$, and cluster our standard errors by village.

$$\begin{aligned} Y_{ij}^{el} =& \beta_0 + \beta_1 T_j + \beta_2 A_{ij}^{bl} + \beta_3 Y_{ij}^{bl} + \beta_4 T_j \times A_{ij}^{bl} \\ &+ \beta_5 T_j \times Y_{ij}^{bl} + \sum_{k=1}^{K} \left[\mu_k X_{ij}^k + \delta_j T_i \times X_{ij}^k \right] + \boldsymbol{S_{ij}}'\gamma + \varepsilon_{ij} \end{aligned} \tag{1}$$

In our preferred specification, we interact T_j with the full set of baseline covariates X_i^k from Table 1. We de-mean A_i^{bl}, Y_i^{bl}, and X_i^k prior to building the interaction terms, so the main effect of the treatment β_1 retains its interpretation as the sample-average treatment effect (Imbens and Rubin 2015). In other robustness checks, we also control for baseline beliefs, Y_i^{bl}, non-linearly.

In Equation 2, we test for the effect of the information treatment on endline imprecision, and heterogeneity in that effect by baseline risk beliefs and imprecision. Here, the outcome variable is endline imprecision A_{ij}^{el}.

$$\begin{aligned} A_{ij}^{el} =& \beta_0 + \beta_1 T_j + \beta_2 A_{ij}^{bl} + \beta_3 Y_{ij}^{bl} + \beta_4 T_j \times A_{ij}^{bl} \\ &+ \beta_5 T_j \times Y_{ij}^{bl} + \sum_{k=1}^{K} \left[\mu_k X_{ij}^k + \delta_j T_i \times X_{ij}^k \right] + \boldsymbol{S_{ij}}'\gamma + \varepsilon_{ij} \end{aligned} \tag{2}$$

All explanatory variables remain the same as in Equation 1 and standard errors remain clustered at the village level.

4 Results

4.1 Characterizing Imprecision

We begin by documenting that imprecise probabilities can be measured separately from levels of risk beliefs. Figure 3 shows the distributions of the level of risk beliefs and imprecision in risk beliefs for both annual (Panel A) and per-act risk beliefs (Panel B). The level of risk beliefs exhibits a large mass point at 100%.[13] People with risk beliefs of 100% and those with lower risk beliefs are largely similar on observed characteristics, other than some differences on cognitive measures and risk attitudes (Table A2). The similarity seen between the reported annual and per-act risks in Figure 3 is unlikely to be driven by scope neglect (Diamond and Hausman 1994, Kamenica 2019) or participants' innumeracy, and instead reflects actual beliefs held by the respondents. Given that the per-act risk is extremely high (Godlonton, Munthali, and Thornton 2016), the similarity in responses can be explained by a simple binomial conversion between per-act and per-year risk beliefs. For example, consider a per-act risk belief of 90% (the median in our sample is 100% and the average is 78%). The average respondent reports 1.5 exposures per week (78 per year), so the per-year belief is $1-(1-0.9)^{78}$ which is nearly 100%. These numbers are even closer for higher-risk beliefs.

The modal person has zero imprecision in our sample, but the distribution has a long tail. The proportion of people with zero imprecision is starkly higher than that in Delavande, Bono, and Holford (2021), where only about 1% of the sample has no imprecision about the risk of catching COVID-19. This large difference might be due to the fact that at the time of Delavande, Bono, and Holford (2021)'s study SARS-CoV-2 was still a novel virus, while our data come from a location and time where HIV was widely well-understood. Reported imprecision around dementia in Giustinelli, Manski, and Molinari (2022*a*) is lower than Delavande, Bono, and Holford (2021), with only about half the sample reporting some imprecision for beliefs around dementia—a disease that people have much more experience with than COVID. The extremely low rates of imprecision in our own sample are likely because HIV is a

[13] This is potentially consistent with the S-shaped probability weighting in Tversky and Kahneman (1992): we do not see people shading down their high probability beliefs as predicted in the original prospect theory paper (Kahneman and Tversky 1979).

well understood disease in Malawi with people across age groups being aware of the disease. (Dementia on the other hand generally affects older people, so people do not start thinking about it until later in life.)[14] Additionally, another study with the same sample shows that extremely high HIV risk beliefs in Malawi cause people to take more risks, i.e. behave fatalistically (Kerwin 2025). Thus, our measure of risk beliefs is likely to be a meaningful measure of people's perceived risks, since it influences people's behavior.

Imprecision has an inverse U-shaped relationship with risk beliefs, consistent with the standard deviation of a Bernoulli (binary) random variable (Figure 4). Individuals at the lowest and highest ends of risk beliefs show the least imprecision, while imprecision is highest for individuals with risk beliefs in the 40-60% range. This pattern, which holds for both annual (Panel A) and per-act risk beliefs (Panel B), illustrates that imprecision is not a mechanical function of the level of risk beliefs. This relationship is underscored by a binned scatterplot of the "upside" and "downside" ranges (the extent to which range goes higher or lower than the point estimate) against the level of risk beliefs (Figure 5).[15] In particular, the downside range (the most by which the respondent thinks that the HIV transmission rate might be below their point estimate) is not constrained at a belief of 100, and can vary to as much as 100 percentage points. Despite this, the downside range is still lower at a belief level of 100 than it is in any other bin besides the very lowest one. Our findings also align with Enke and Graeber (2023)'s result that subjective probabilities associated with higher "cognitive uncertainty" are compressed towards 50:50. We are able to rule out the possibility that this is partly driven by mechanical effects of being at the edge of the parameter space, which Enke and Graeber (2023) cannot.

We also see that imprecision is negatively correlated with risk beliefs throughout the length of the study—starting from the baseline to the endline three months after the information treatment. Figure 2 presents the mean annual risk beliefs for treatment-group respondents at three points in time: the initial survey question, an immediate followup question just after the information treatment, and the endline

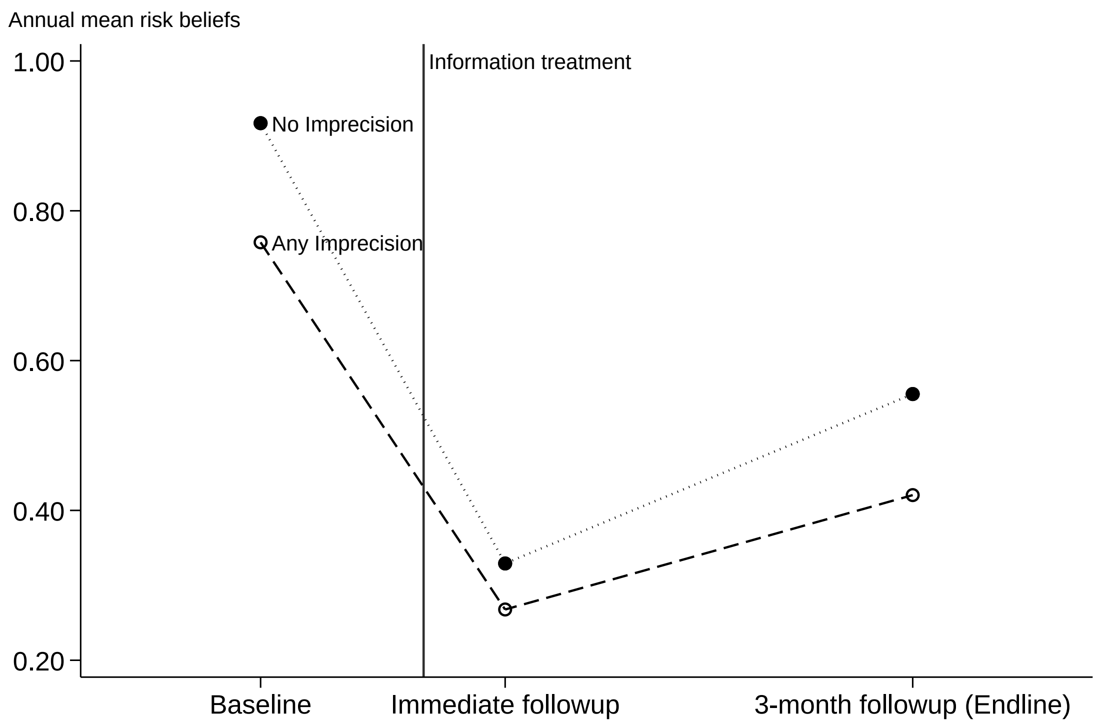
[14] In a very different context, Bachmann et al. (2020) find similar imprecision to our study: about 20% of firm-quarter observations report non-zero imprecision when asked about the probability of an increase in future sales in the current quarter.

[15] The upside and downside ranges capture the distance between question H1a and questions H1c and H1d, respectively, in the survey questions shown in Figure A1.

survey. The line for respondents who express any imprecision is consistently lower than the one for those with no imprecision. The figure also shows that the effect of the information treatment appears to weaken over time. This is consistent with respondents receiving other signals about the risk of HIV transmission over time (such as hearsay from friends) that are closer to their original priors than to the content of the information treatment.

Figure 2

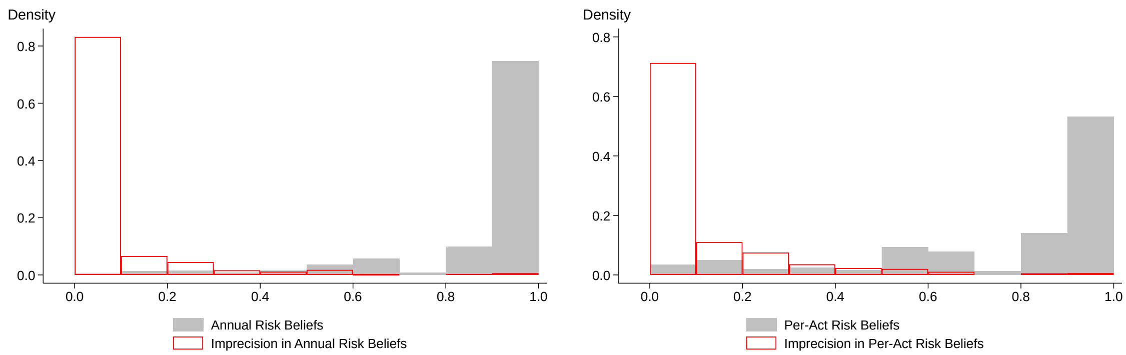
Risk Beliefs by Imprecision and Survey Wave



Notes: Sample is 647 people who completed both a baseline and an endline survey and were in the information treatment group. The solid vertical line indicates the timing of the information treatment during the baseline survey. Black circles denote the mean risk beliefs of people with no imprecision in their risk beliefs and hollow circles denote the mean risk beliefs of people with imprecision.

Imprecision is also related to the tendency to report "rounded probabilities" in their risk beliefs, such as values divisible by 0.05 or 0.50. Our preferred specifications in columns 6 and 7 in Table 3 respectively control for whether the respondent's original answer to question H1a and the probing question H1b in Figure A1 are divisible by 0.5. These results show that on average individuals with rounded beliefs display imprecision that is about 0.04 percentage points higher, which is 80% higher than the mean. This is consistent with Giustinelli, Manski, and Molinari (2022*a*), who note that survey expectations are generally rounded, and find that respondents who report rounded beliefs are often willing to give a range for their beliefs.

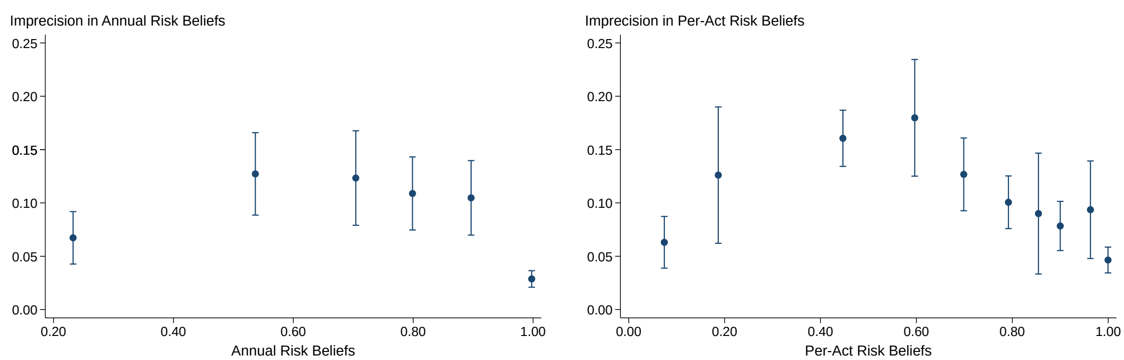
Figure 3
Distribution of Risk Beliefs and Imprecision



Panel A: Annual Risk Beliefs　　**Panel B:** Per-Act Risk Beliefs

Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey. Both variables are measured at baseline. Annual risk beliefs are the perceived chance of contracting HIV from regular unprotected sex with an HIV-positive sexual partner over a period of one year. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk.

Figure 4
Imprecision by Level of Risk Beliefs

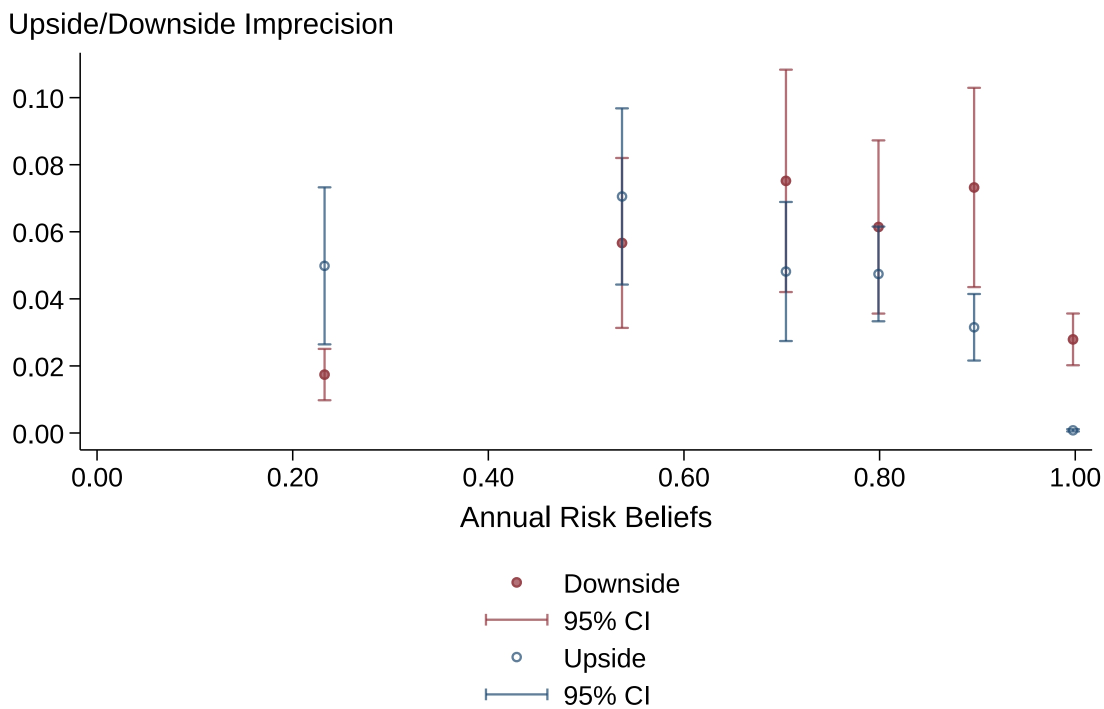


Panel A: Annual　　**Panel B:** Per-Act

Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey. Both variables are measured at baseline. Annual risk beliefs are the perceived chance of contracting HIV from regular unprotected sex with an HIV-positive sexual partner over a period of one year. Per-act risk beliefs are the perceived chance of contracting HIV from a single unprotected sex act with an HIV-positive sexual partner. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk.

Figure 5

Upside and Downside Imprecision by Level of Risk Beliefs



Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey. Both variables are measured at baseline. Annual risk beliefs are the perceived chance of contracting HIV from regular unprotected sex with an HIV-positive sexual partner over a period of one year. Upside Imprecision is the distance between the respondent's point estimate of the risk and the highest value they think the risk could take. Downside Imprecision is the distance between the respondent's point estimate of the risk and the lowest value they think the risk could take.

Table 3

Relationship between Imprecision and Rounded Beliefs

	Outcome: Imprecision in Annual Risk						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Annual HIV transmission risk divisible by 0.05	-0.013		-0.014		0.037**	0.037**	0.038**
	(0.015)		(0.016)		(0.018)	(0.018)	(0.018)
Annual HIV transmission risk divisible by 0.50		-0.063***		-0.064***	-0.069***	-0.069***	-0.073***
		(0.009)		(0.009)	(0.010)	(0.010)	(0.010)
Annual HIV transmission risk originally divisible by 0.50						0.004	
						(0.061)	
Thinks answer is really 0.50							0.088***
							(0.031)
Controls for Baseline Covariates	No	No	Yes	Yes	Yes	Yes	Yes
Observations	1,284	1,284	1,237	1,237	1,237	1,237	1,237
Adjusted R-squared	0.004	0.047	0.002	0.046	0.048	0.047	0.060
Control-group Mean	0.041	0.041	0.040	0.040	0.040	0.040	0.040
Control-group SD	0.105	0.105	0.105	0.105	0.105	0.105	0.105

Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk. Other baseline covariates include all variables in Table 1. All regressions control for stratification cell fixed effects. Heteroskedasticity-robust standard errors, clustered by village, in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. "Annual HIV transmission risk originally divisible by 0.5" corresponds to only the originally stated risk beliefs in question H1a being divisible by 0.5 and "Thinks answer is really 0.50" corresponds to the probing question H1b in Appendix Figure A1.

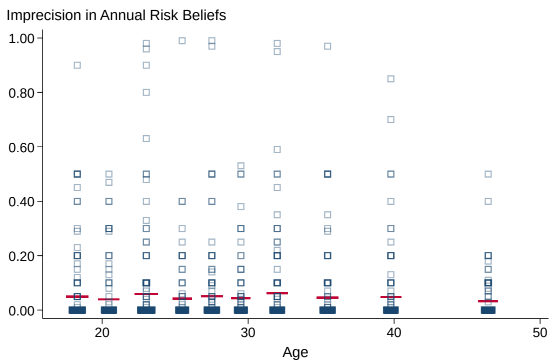
There are no strong relationships between imprecision and observable characteristics. Appendix Figure A2 shows a histogram of imprecision in annual risk beliefs (our preferred measure) by gender. We see that imprecision follows a similar distribution for both females and males, with less than 20% of both genders showing any imprecision at all. Furthermore, we find no strong patterns of imprecision by participants' observable characteristics such as age, education, numeracy, and years sexually active.

The lack of a relationship between imprecision in risk beliefs and education is surprising, but we believe that it does not reflect a lack of understanding of subjective probabilities in our study sample. The average respondent in our sample has only about six years of schooling; as noted by Delavande, Giné, and McKenzie 2011, however, developing-country study participants with limited education can understand and respond to questions on subjective probabilities. Among our study participants, numeracy—as expected—is strongly correlated with education, and baseline risk beliefs are higher in respondents with more education (Figure A3). Relatedly, we also see that people who score higher on a cognitive index have higher baseline risk beliefs, but we do not see such an association with imprecision (Figure A4), suggesting a true lack of correlation between cognitive abilities and imprecision. Furthermore, we do not see a significant correlation between baseline sexual activity and baseline imprecision (Table A17). Finally, following Delavande and Kohler (2009) we also compare the standard deviation of risk beliefs across different outcomes to assess whether respondents might be answering questions at random. For example, the standard deviations of imprecision in annual and per-act HIV transmission beliefs are 0.13 and 0.16 respectively, and we can reject the null hypothesis that they are equal ($F = 0.72$, cluster-adjusted $p = 0.003$). This suggests that the responses are not random, and instead reflect meaningful differences across variables.

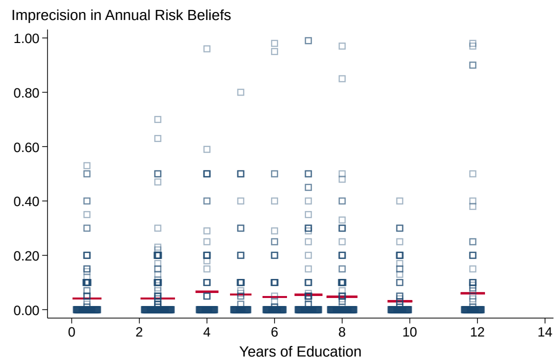
Figure 6 presents strip plots of imprecision by individual characteristics. These have a similar interpretation to histograms: each individual data point is shown as a hollow square, and the plot is both wider and denser when there is more data at a specific value. The means for each category are indicated with red lines. These plots show that both the mean and the overall distribution of imprecision are very similar across all levels of the observed covariates. We also note that endline imprecision is

Figure 6

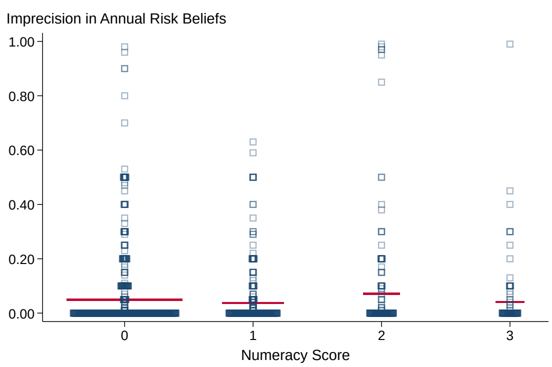
Strip Plots of Imprecision vs. Baseline Covariates (Annual Risk Beliefs)



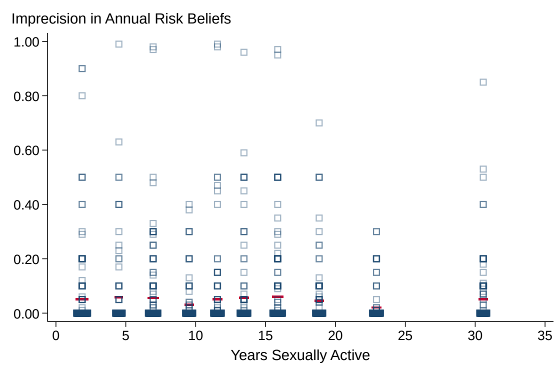
Panel A: Imprecision vs. Age



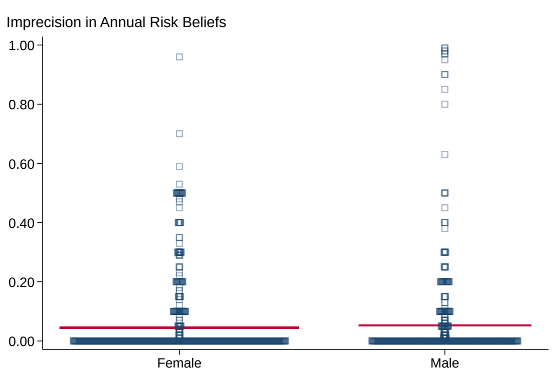
Panel B: Imprecision vs. Education



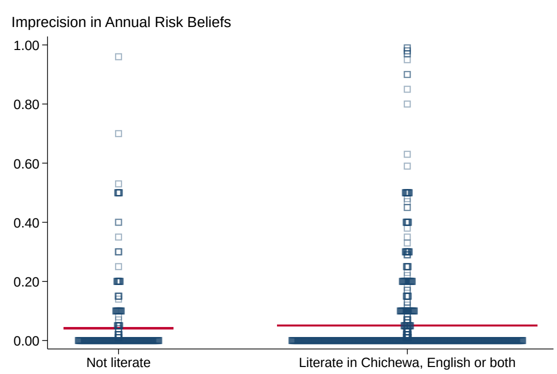
Panel C: Imprecision vs. Numeracy



Panel D: Imprecision vs. Years Sexually Active



Panel E: Imprecision vs. Gender



Panel F: Imprecision vs. Literacy

Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey. Both variables are measured at baseline. Annual risk beliefs are the perceived chance of contracting HIV from regular unprotected sex with an HIV-positive sexual partner over a period of one year. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk. Each hollow square is an individual data point.

significantly but weakly predicted by baseline imprecision (Table A18).

We test the predictive power of all the baseline covariates in Table 1, and find little evidence that they predict either annual or per-act imprecision (see Appendix Table A3 and Table A4). For imprecision in annual risk beliefs, neither bivariate nor multivariable regressions show any significant correlations with baseline covariates. We also try two machine-learning methods: kernel-regularized least squares (KRLS, Hainmueller and Hazlett 2014) and the LASSO (Friedman, Hastie, and Tibshirani 2010). KRLS shrinks the effects of all covariates nearly to zero, and none are statistically significant. The LASSO is designed to shrink small coefficients to zero, so it is not surprising that it does not pick any variables; this empirical pattern mirrors the finding of Cilliers, Elashmawy, and McKenzie (2024) that the post-double-selection LASSO rarely selects many variables in RCTs. In line with those results, the R^2 values for the multivariable regression, KRLS, and the LASSO are all less than 0.005. For per-act risk imprecision, "Risk Attitude" and "Lifetime Sex Partners" show some predictive power across the different estimators, but the overall R^2 value remains very low. We also test the power of covariates for predicting a non-zero range among the respondents in our sample and find limited evidence of predictive power (Appendix Table A5 and Table A6). Conditional on other covariates, men are more likely to express any imprecision in their beliefs, as are people from larger households and those who are inclined to take more risks.

A potential concern is that the high percentage of zeros in our data could limit our ability to estimate between imprecision and covariates. To address the presence of the large number of zeroes in our data, we test the power of covariates for predicting a non-zero range among the respondents in our sample and find limited evidence of predictive power (Appendix Table A5 and Table A6). Conditional on other covariates, men are more likely to express any imprecision in their beliefs, as are people from larger households and those who are inclined to take more risks. To account for the large number of zeros in the imprecision measure, we also conduct a Tobit analysis (Table A7), which again finds limited evidence of predictive power for both annual and per-act beliefs. Moreover, it is likely that individuals' certainty in their priors does not vary much with the demographic and other characteristics in our sample. Using multinomial probit, Giustinelli, Manski, and Molinari (2022*a*) in their dementia risk beliefs paper, do find significant associations of imprecision with age and race. But they do not find significant associations with higher education, gender, and cognitive ability. Delavande, Bono, and Holford (2021) use an OLS regression and see some differences in the imprecision of COVID infection and transmission beliefs by respondent's race and GPA, and find no significant differences by gender, parent's education, and residency status. Finally, Bachmann et al. (2020) find imprecision to be more common among small firms. Thus, our results generally align with those of other papers, although the correlations we find with covariates are somewhat smaller.

4.2 The Role of Imprecision in Risk Belief Updating

Next, we examine how imprecision influences the updating of risk beliefs, and compare our findings with the predictions from our two models. Panel A of Table 4 shows the results of estimating Equation 1. The effect of the information treatment on how people update their beliefs depends on their baseline level imprecision; those with higher imprecision levels update more in response to the treatment. In our preferred specification in column 3, we see that for every 10 percentage point increase in imprecision, the magnitude of the effect of the treatment increases by about 2.5 percentage points. These results hold even after controlling for the full set of baseline characteristics from Table 1 and their interactions with the treatment indicator (column 4).

The finding that the updating of risk beliefs depends on people's imprecision levels holds only when we control for the interaction between baseline risk beliefs and the treatment. This makes sense because of the strong correlation between the two variables. As shown in Figure 4, however, this relationship is non-linear. Thus, in Table 5 we show that the results for annual beliefs, which is our preferred outcome variable, remain robust to controlling for the level of risk beliefs non-linearly. Specifically, we control for indicators for bins of baseline risk beliefs and their interactions with the treatment indicator (see Appendix Table A8 for the ranges of the baseline beliefs for each of the bins). In column 4, which also controls for baseline covariates and their interactions with the treatment indicator, we find that a 10 percentage-point increase in imprecision increases the magnitude of the treatment effect by 2.1 percentage points. This is a substantively large effect: it means that people at the

90$^{\text{th}}$ percentile of imprecision (with a range of 20 percentage points) will update their beliefs by an additional 4.3 percentage points relative to someone at the median (with a range of 0 percentage points). Given that the average treatment effect on risk beliefs is 36 percentage points, this is an increase in the treatment effect of more than 10%.

Panel A of Appendix Figure A5 shows the estimated effects by bins of baseline risk beliefs from the same regression.[16] People at the lowest end of the risk belief spectrum have the smallest treatment effects while those at the middle and highest ends have effects that are relatively larger in magnitude.[17] These results help justify the need for controlling for the interactions between treatment and baseline beliefs in our regressions: not only is imprecision correlated with the levels of beliefs, but the level of beliefs is correlated with treatment effects, creating a potential source of omitted variable bias.

The limited correlations between imprecision and respondent characteristics noted in Section 4.1 are advantageous for our analysis. A key concern for our findings is that any relationship between imprecision and an outcome of interest could instead be driven by omitted variables that are correlated with both imprecision and the outcome. The same concern applies to our analyses of treatment effect heterogeneity: it is difficult to rule out the possibility that the treatment effects actually vary by some other factor that is correlated with imprecision. Our primary way of addressing this concern is to control for both a vector of other respondent characteristics and their interactions with the treatment indicator. However, if these characteristics are measured with error, then controlling for them may not completely address the confounding problem (Pei, Pischke, and Schwandt 2019). In that situation, Pei, Pischke, and Schwandt recommend showing balance on these other covariates; the strip plots in Figure 6 are a version of their test.

[16] Each bin is shown along the x-axis at the average value of risk beliefs for that bin, so e.g. the bottom bin is shown at 0.4.

[17] We can also include the three-way interaction between imprecision, brackets of risk beliefs, and treatment status (Appendix Table A9). The three-way interactions are quite noisily estimated, but there is suggestive evidence that the heterogeneity by imprecision is driven by people in the middle of the risk belief distribution.

Table 4

Treatment Effect Heterogeneity by Imprecision

	Outcome: Endline Annual HIV Transmission Risk				*Outcome: Endline Per-Act HIV Transmission Risk*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment Group	-0.371***	-0.369***	-0.367***	-0.358***	-0.384***	-0.381***	-0.379***	-0.374***
	(0.016)	(0.016)	(0.017)	(0.016)	(0.019)	(0.019)	(0.019)	(0.018)
Imprecision		-0.069	0.008	0.024		-0.150*	-0.070	-0.092
		(0.072)	(0.079)	(0.076)		(0.087)	(0.092)	(0.085)
Baseline Beliefs	0.274***	0.254***	0.501***	0.480***	0.272***	0.262***	0.447***	0.428***
	(0.048)	(0.048)	(0.061)	(0.058)	(0.033)	(0.034)	(0.047)	(0.052)
Treatment × Imprecision		-0.184	-0.291**	-0.246**		0.020	-0.083	-0.073
		(0.117)	(0.119)	(0.113)		(0.111)	(0.116)	(0.119)
Treatment × Baseline Beliefs			-0.379***	-0.349***			-0.301***	-0.282***
			(0.081)	(0.080)			(0.060)	(0.067)
T Interacted w/ Other Baseline Covariates	No	No	No	Yes	No	No	No	Yes
Observations	1,276	1,276	1,276	1,230	1,281	1,279	1,279	1,232
Adjusted R-squared	0.328	0.333	0.342	0.394	0.315	0.318	0.329	0.355
Control-group Mean	0.906	0.906	0.906	0.907	0.743	0.743	0.743	0.742
Control-group SD	0.196	0.196	0.196	0.195	0.317	0.317	0.317	0.317

Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk. Other baseline covariates include all variables in Table 1. All regressions control for stratification cell fixed effects. Heteroskedasticity-robust standard errors, clustered by village, in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 5

Treatment Effect Heterogeneity by Imprecision
(Adjusting for Heterogeneity by Brackets of Beliefs)

	Outcome: Endline Annual HIV Transmission			
	(1)	(2)	(3)	(4)
Treatment ×				
Bottom Bin of Baseline Beliefs	-0.173***	-0.158***	-0.174***	-0.158***
	(0.057)	(0.057)	(0.057)	(0.057)
Middle Bin of Baseline Beliefs	-0.425***	-0.411***	-0.417***	-0.405***
	(0.051)	(0.053)	(0.054)	(0.056)
Top Bin of Baseline Beliefs	-0.391***	-0.392***	-0.378***	-0.381***
	(0.018)	(0.018)	(0.017)	(0.017)
Imprecision		-0.020		0.002
		(0.081)		(0.079)
Treatment × Imprecision		-0.250**		-0.213*
		(0.119)		(0.114)
T Interacted w/ Other Baseline Covariates	No	No	Yes	Yes
Observations	1,276	1,276	1,230	1,230
Adjusted R-squared	0.334	0.339	0.389	0.389
Control-group Means				
Overall	0.906	0.906	0.907	0.907
Bottom Bin	0.643	0.643	0.640	0.640
Middle Bin	0.847	0.847	0.842	0.842
Top Bin	0.937	0.937	0.938	0.938
Control-group SD	0.196	0.196	0.195	0.195

Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk. Other baseline covariates include all variables in Table 1. All regressions control for stratification cell fixed effects. Heteroskedasticity-robust standard errors, clustered by village, in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

The lack of a correlation between imprecision and other respondent characteristics makes it extremely unlikely that confounding (by those characteristics) is driving our results. A potential downside of the lack of correlation between imprecision and respondent characteristics is that it could indicate that our imprecision measure is not capturing a real signal about people's beliefs. Delavande, Giné, and McKenzie (2011) argue that one reason to believe that subjective expectation measures work well in developing countries is that they correlate with other characteristics. However, our other results provide evidence that imprecision is in fact capturing a real feature of respondents' beliefs. In particular, we find that imprecision is not merely a mechanical function of the level of risk beliefs, and also that it plays a role in belief updating.

Our findings for belief updating match what our standard Bayesian model predicts: people with more imprecision update their beliefs by more. This is intuitive, since these are the people who are less sure about their priors and thus are more persuadable. These results differ, however, from the predictions of the Robust Bayes model, in which there are no differences in updating by prior imprecision levels. We believe that our findings offer evidence of actual learning in the face of new information, similar to that in Dupas (2011), which shows that the provision of credible information on HIV risk led to sexual behavior change among teenagers.

4.3 The Effects of New Information on Imprecision

The information treatment also changes respondents' imprecision about their beliefs. However, contrary to the predictions of both versions of the model, we see that the effect varies with the levels of risk beliefs that individuals hold. In Table 6 we see that the average treatment effect on endline imprecision is negligible and statistically insignificant for both annual (Columns 1-4) and per-act risk beliefs (Columns 5-8). However, for annual imprecision, this effect varies sharply by baseline risk beliefs (Column 4): people with lower priors have negative treatment effects (so the information treatment makes their beliefs more precise), and those with higher priors have positive ones (so the treatment makes their beliefs less precise).

To explore this further, we control for the levels of beliefs non-linearly in Table 7. This does not affect the heterogeneity in treatment effects by imprecision, which remains statistically insignificant. It does reveal that the heterogeneity by baseline beliefs is clearly non-linear. Panel A of Appendix Figure A6 plots the treatment effect for each bin of baseline risk beliefs. Again, our findings deviate from both models but the results are consistent with what we would expect given the nature of the shock to risk beliefs in our experiment. Respondents at the lowest end of the belief spectrum face the smallest shock to priors about HIV transmission risk, since the information treatment told them that they were fairly close to being correct already. They thus show a negligible treatment effect. In contrast, respondents at the highest end of the risk belief spectrum on the other hand face the largest shock to their priors, and thus increase their imprecision.

While these results are intuitive, they contrast with the theoretical predictions in Section 1. Both versions of the model predict that updating of imprecision should be independent of the prior mean of risk beliefs, and thus that we should not observe any heterogeneity in treatment effects by baseline risk beliefs. This implies that these models of belief updating are misspecified in some way. One possibility for improving the predictive performance of the model is to use a different prior for risk beliefs, rather than the Beta distribution. Alternatively, it may be necessary to change the model of updating entirely.

4.4 Robustness checks

We subject our results to several robustness checks. First, as discussed above, our main results are robust to controlling for the interactions between baseline characteristics and the treatment indicator, and to controlling for the levels of risk beliefs non-linearly. Second, our results are robust to dropping the highest and the lowest risk beliefs (Appendix Table A10), thereby showing that the results are not driven by people with extremely high or extremely low beliefs about HIV transmission. This is also true if we control for the level of risk beliefs using brackets for ranges of beliefs (Appendix Table A11). We also find the same qualitative pattern if we restrict the sample to respondents with non-zero imprecision (Appendix Table A12 and Appendix Table A13). This sample has only 294 observations, so these results are noisily estimated, and are significant for annual risk beliefs only at the $\alpha = 0.10$ level.

Table 6

Treatment Effects on Imprecision

	Outcome: Endline Imprecision in Annual Risk				*Outcome: Endline Imprecision in Per-Act Risk*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment Group	0.011	0.008	0.008	0.010	-0.012	-0.014*	-0.014*	-0.010
	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)	(0.009)
Imprecision		0.225***	0.208***	0.197***		0.146**	0.145**	0.148**
		(0.070)	(0.070)	(0.070)		(0.056)	(0.058)	(0.062)
Baseline Beliefs	-0.017	-0.000	-0.055*	-0.043	-0.001	0.008	0.006	0.013
	(0.017)	(0.016)	(0.031)	(0.036)	(0.014)	(0.015)	(0.024)	(0.025)
Treatment × Imprecision		-0.139	-0.116	-0.087		-0.077	-0.076	-0.088
		(0.085)	(0.086)	(0.089)		(0.068)	(0.072)	(0.074)
Treatment × Baseline Beliefs			0.084**	0.087**			0.003	0.001
			(0.036)	(0.041)			(0.029)	(0.032)
T Interacted w/ Other Baseline Covariates	No	No	No	Yes	No	No	No	Yes
Observations	1,270	1,270	1,270	1,224	1,276	1,274	1,274	1,228
Adjusted R-squared	0.006	0.035	0.039	0.051	-0.001	0.010	0.009	0.009
Control-group Mean	0.041	0.041	0.041	0.040	0.041	0.074	0.074	0.073
Control-group SD	0.105	0.105	0.105	0.105	0.143	0.144	0.144	0.142

Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk. Other baseline covariates include all variables in Table 1. All regressions control for stratification cell fixed effects. Heteroskedasticity-robust standard errors, clustered by village, in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 7
Treatment Effects on Imprecision
(Adjusting for Heterogeneity by Brackets of Beliefs)

	Outcome: Endline Imprecision in Annual Risk			
	(1)	(2)	(3)	(4)
Treatment ×				
Bottom Bin of Baseline Beliefs	-0.008	-0.007	-0.005	-0.006
	(0.014)	(0.015)	(0.017)	(0.017)
Middle Bin of Baseline Beliefs	-0.072***	-0.064**	-0.073***	-0.065**
	(0.026)	(0.026)	(0.027)	(0.026)
Top Bin of Baseline Beliefs	0.024***	0.020***	0.025***	0.023***
	(0.007)	(0.007)	(0.008)	(0.008)
Imprecision		0.177**		0.158**
		(0.068)		(0.067)
Treatment × Imprecision		-0.084		-0.047
		(0.084)		(0.087)
T Interacted w/ Other Baseline Covariates	No	No	Yes	Yes
Observations	1,270	1,270	1,224	1,224
Adjusted R-squared	0.033	0.053	0.050	0.071
Control-group Means				
Overall	0.041	0.041	0.040	0.040
Bottom Bin	0.049	0.049	0.046	0.046
Middle Bin	0.125	0.125	0.130	0.130
Top Bin	0.030	0.030	0.030	0.030
Control-group SD	0.105	0.105	0.105	0.105

Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk. Other baseline covariates include all variables in Table 1. All regressions control for stratification cell fixed effects. Heteroskedasticity-robust standard errors, clustered by village, in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Third, our core results are unaffected by potential alternative ways of constructing our measure of imprecision. One potential issue with using the full range that respondents report is that the range is bounded mechanically on one side for certain risk beliefs: if the level of risk beliefs is already 1, it is not possible to report a higher upper bound; a similar issue applies to beliefs close to 1 (or close to 0). To address this we construct a "half-width", which is the gap between the point estimate and whichever of questions H1c and H1d in Figure A1 is farthest away from the point estimate. By capturing the wider of the sub-ranges that extend above and below the level of risk beliefs, this half-width measure provides an alternative measure of imprecision that is not bounded above at 100% for respondents with high risk beliefs, nor below at 0% for respondents with low risk beliefs. This half-width measure, when plotted against the level of risk beliefs, follows the same inverse U-shaped pattern as our original measure (Appendix Figure A7).[18] Table A14 shows treatment effects on endline beliefs using this alternate definition for both annual and per-act risk beliefs. Our results for annual risk beliefs remain robust to the use of this alternate definition, with the treatment effect increasing by about 3 percentage points for every 10 percentage-point increase in imprecision.

This half-width measure introduces another potential issue: if people's risk beliefs are at 0.5, their half-width can be at most 0.5 (extending up to 1 or down to 0). On the other hand, in principle someone could report a belief level of 1 with a range extending down to 0, leading to a half-width of 1—and this does happen for some observations in our sample. To correct for this disparity, we thus construct a truncated half-width that top-codes the original half-width measure at 50 percentage points. Similar to the original half-width measure, this truncated measure exhibits an inverse U-shaped relationship with baseline beliefs (Appendix Figure A8).[19] In Table A15 we show that the belief updating results for annual risk beliefs also remain robust to this top-coding.

Another approach is to collapse imprecision into an indicator that is equal to one if there is any imprecision and zero otherwise. Doing this yields similar results to the continuous measurement (Appendix Table A16), but is somewhat noisily estimated, with $p = 0.103$ in our preferred specification in Column 3. A related issue is that people may under-report their imprecision because they are ambiguity averse, and do not want to admit that they are unsure about the probabilities in question. In this case, some of the people who report ranges of zero actually have non-zero imprecision about their beliefs. We conduct a simple Monte Carlo simulation to test how this would affect our results by taking our actual data and randomly setting the range

[18] Panel A of Appendix Figure A9 plots the difference between the half-width and the original range by the level of risk beliefs.

[19] Panel B of Appendix Figure A9 plots the difference between the truncated half-width and the original range by the level of risk beliefs.

to zero with 50% probability. While this measurement error is not classical, we find that on average it slightly attenuates the estimated coefficient on the interaction term and increases the p-values. Thus this sort of measurement error would mean that our results understate the role of imprecision in belief updating.

Fourth, we explore an alternate measure of risk beliefs: per-act, rather than annual, risks. The basic descriptive facts about imprecision also hold for this alternate measure: they have an inverse U-shaped relationship with baseline risk beliefs and are not strongly predicted by other covariates. However, as shown in Columns 5-8 of Table 4 and Table 6 and the corresponding figures in Panels B of Figure A5 and Figure A6, our results for belief updating and effects on imprecision do not hold for per-act risk beliefs. The one exception to this is if we collapse imprecision into a binary variable (Appendix Table A16, Columns 5-8). The weaker evidence of an effect for per-act beliefs may be due to the fact that the information treatment told respondents about the true *annual* risk of HIV transmission. This information is informative about what the per-act risks must be, and we do see respondents updating their risk belief levels in response to the information. The weak evidence of updating for per-act risks is also suggestive of how participants form their priors: they are provided with information about annual risks and they work out per-act risks from this information. However, it is less clear exactly how imprecision should be related to this updating process, and so the pattern (if any) may be less stark and harder to detect in our data. We nevertheless take this result as a potential caveat to our findings.

4.5 Effects on Endline Sexual Activity

The information treatment increased sexual activity among the participants by about 10-12% on average (Appendix Table A19), which is consistent with conventional risk compensation. Kerwin (2025) shows that this pattern varies sharply by the level of baseline risk beliefs: people with high priors substantially *reduce* how much sex they have, because the information treatment reduces fatalism. We examine how these effects vary by baseline imprecision levels. For people with higher imprecision, the information treatment has a less positive impact on sexual activity, but these estimates are noisy and we do not come close to rejecting the null hypothesis that treatment effects are uncorrelated with imprecision. The confidence intervals are wide, so these null findings are not particularly informative. Our preferred specification is shown in column 4 (for annual risk beliefs) and column 10 (for per-act risk beliefs). In these specifications, a 100 percentage-point increase in imprecision is associated with a 30-40 log point reduction in treatment effects, relative to an average treatment effect of 10 log points. The confidence interval around these estimates ranges from a 110 log point reduction to a 30 log point increase. If we took the point estimates literally, they would imply that information campaigns should be targeted on the basis of people's information levels—since their effectiveness could vary widely based on how sure people are of their baseline beliefs.

5 Conclusion

Using experimental data from Malawi, we show that people express uncertainty about their risk beliefs that can be measured separately from the levels of those beliefs. We measure these "imprecise probabilities" by asking people to express a range around their beliefs about HIV transmission risks. Imprecision in risk beliefs covaries with the levels of risk beliefs exactly as we would expect, with higher values in the middle of the belief distribution. At the same time, it is not perfectly determined by belief levels, and plays a separate role in belief updating. In this context—where people's beliefs about HIV risk transmission are much higher than the truth—providing people with information about the true risk of HIV transmission lowers people's risk beliefs. This updating pattern is stronger for those with higher levels of imprecision. The provision of new information also shifts people's imprecision levels: people with low priors (whose beliefs are reinforced) reduce their imprecision, while those with high priors increase their imprecision. These results are consistent with some, but not all, of the predictions of Bayesian models of belief updating. They are a better match for our classical Bayesian model than the Robust Bayes approach which is designed to handle imprecise probabilities. However, neither version of the model is a satisfactory match for all of our estimates. Moreover, both models also deviate from the intuition that surprising information should (all else equal) raise imprecision. Further work should develop models that can capture all aspects of the belief updating process.

A key limitation of our belief updating results is that they depend on the measure of risk beliefs we use. We find the same basic patterns of imprecision for both per-act and annual HIV risk beliefs. However, the pattern of treatment effect heterogeneity by baseline imprecision, and the changes in imprecision in response to the information treatment, hold only for annual risk beliefs. One explanation for this difference is that the information treatment itself is about annual risk beliefs. While this also provides information about the implied per-act risk (and people do update the levels of their per-act risk beliefs), the role of imprecision in the belief updating process is clearer for annual than for per-act beliefs.

The role of imprecision in belief updating is relevant for the design of information campaigns. Imprecision in risk beliefs is a measure of how rigid and immovable people's beliefs are: for the exact same level of risk beliefs, people can vary in their imprecision, and thus in their willingness to change their minds. The same information campaign, targeted at people who show more imprecision, will lead to larger changes in risk beliefs—and thus potentially to larger behavioral changes as well. For instance, an intervention aimed at encouraging a given population to sign up for age-specific health screenings in order to reduce mortality from breast or prostate cancer, might be more successful if people's priors are more malleable. This is consistent with previous evidence from Han et al. (2007), who find that imprecision in beliefs about cancer prevention recommendations influences people's uptake of cancer screening. It is also related to the literature on Bayesian persuasion (Kamenica 2019).

Moreover, information campaigns might also be more effective if they targeted imprecision levels directly. Since people with higher levels of imprecision update their beliefs more, increasing people's imprecision could make them more willing to change their priors. Consider, for example, people's beliefs about the effectiveness of the COVID-19 vaccine. Rather than simply trying to change these beliefs directly, public health efforts might first raise doubts about the information that led people to form these beliefs in the first place. If people become more unsure of their beliefs about the vaccine's effectiveness, this might make them more receptive to messages communicating the true effectiveness. Increasing people's imprecision about their beliefs could make them more willing to change their minds (and in some cases, it might also leave them susceptible to disinformation). Whether this approach works better than simply providing all the information at once depends on the specifics of the belief updating process. Future research should explore this updating process in more detail.

Understanding why people with the exact same risk beliefs differ in their imprecision about those beliefs remains an open question. Future work should seek to understand the factors that shape imprecise probabilities. Another open question is how context shapes these findings: do people update their beliefs similarly in health and non-health settings? For example, policymakers might want to change people's beliefs about climate change. Unlike HIV, cancer, or COVID-19, climate change's impacts are less personal, and instead are diffused across the entire world. People might therefore update their beliefs about it differently (see Sunstein et al. (2016)). Moreover, our findings shed light on imprecision in risk beliefs in the specific context a HIV—a disease that people in Malawi are well aware of. The prevalence of imprecision as well as the role that imprecision plays in updating of beliefs can vary greatly with different risk contexts such as diseases, climate, and stocks along with people's understanding of these risks. The importance of imprecision with respect to such a widely understood disease implies that it plays a very general role in the formation and updating of risk beliefs. Finally, policy-oriented research should also explore when reducing people's imprecision is helpful, and when it is undesirable. How malleable people's priors are will depend on the specific context and the relative reliability of previous data versus new information.

References

Augenblick, Ned, Eben Lazarus, and Michael Thaler. 2023. "Overinference from Weak Signals and Underinference from Strong Signals." arXiv:2109.09871 [econ, q-fin].

Aydogan, Ilke, Loïc Berger, and Valentina Bosetti. 2023. "Unraveling Ambiguity Aversion." *Review of Economics and Statistics*, in press.

Bachmann, Rüdiger, Stefan Lautenbacher, Kai Carstensen, and Martin Schneider. 2020. "Uncertainty is More Than Risk – Survey Evidence on Knightian and Bayesian Firms." Draft.

Baliga, Sandeep, and Tomas Sjöström. 2008. "Strategic Ambiguity and Arms Proliferation." *Journal of Political Economy*, 116(6): 1023–1057.

Camerer, Colin, and Martin Weber. 1992. "Recent Developments in Modeling Preferences: Uncertainty and Ambiguity." *Journal of Risk and Uncertainty*, 5(4): 325–370.

Cameron, Trudy Ann. 2005. "Updating Subjective Risks in the Presence of Conflicting Information: An Application to Climate Change." *Journal of Risk and Uncertainty*, 30(1): 63–97.

CDC. 2019. "HIV Risk Behaviors." https://www.cdc.gov/hiv/risk/estimates/riskbehaviors.html.

Charles, Constantin, Cary Frydman, and Mete Kilic. 2022. "Insensitive Investors." *SSRN Electronic Journal*.

Cilliers, Jacobus, Nour Elashmawy, and David McKenzie. 2024. *Using Post-Double Selection Lasso in Field Experiments.* Washington, DC: World Bank.

Delavande, Adeline. 2008. "Measuring Revisions to Subjective Expectations." *Journal of Risk and Uncertainty*, 36(1): 43–82.

Delavande, Adeline. 2014. "Probabilistic Expectations in Developing Countries." *Annual Review of Economics*, 6(1): 1–20.

Delavande, Adeline. 2022. "Expectations in Development Economics." In *Handbook of Economic Expectations.* Frisco, CO:Elsevier.

Delavande, Adeline, and Hans-Peter Kohler. 2009. "Subjective Expectations in the Context of HIV/AIDS in Malawi." *Demographic Research*, 20: 817–874.

Delavande, Adeline, Emilia Del Bono, and Angus Holford. 2021. "Perceived Ambiguity About COVID-Related Health Outcomes and Protective Health Behaviors Among Young Adults." Working Paper.

Delavande, Adeline, Xavier Giné, and David McKenzie. 2011. "Measuring Subjective Expectations in Developing Countries: A Critical Review and New Evidence." *Journal of Development Economics*, 94(2): 151–163.

Diamond, Peter A, and Jerry A Hausman. 1994. "Contingent Valuation: Is Some Number Better than No Number?" *Journal of Economic Perspectives*, 8(4): 45–64.

Dominitz, Jeff. 1998. “Earnings Expectations, Revisions, and Realizations.” *The Review of Economics and Statistics*, 80(3): 374–388.

Dominitz, Jeff, and Charles F. Manski. 1994. “Eliciting Student Expectations of the Returns to Schooling.” National Bureau of Economic Research Working Paper 4936.

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

Eichberger, Jürgen, and David Kelsey. 2002. “Strategic Complements, Substitutes, and Ambiguity: The Implications for Public Goods.” *Journal of Economic Theory*, 106(2): 436–466.

Enke, Benjamin, and Thomas Graeber. 2023. “Cognitive Uncertainty.” *The Quarterly Journal of Economics*, 138(4): 2021–2067.

Ferrari, Silvia L P, and Francisco Cribari-Neto. 2004. “Beta Regression for Modeling Rates and Proportions.” *Journal of Applied Statistics*, 31(7): 799–815.

Fischhoff, Baruch, and Wändi Bruine De Bruin. 1999. “Fifty–Fifty=50%?” *Journal of Behavioral Decision Making*, 12(2): 149–163.

Friedman, Jerome, Trevor Hastie, and Rob Tibshirani. 2010. “Regularization Paths for Generalized Linear Models via Coordinate Descent.” *Journal of Statistical Software*, 33(1): 1–22.

Gärdenfors, Peter, and Sahlin, Nils-Eric. 1982. “Unreliable Probabilities, Risk-taking, and Decision Making.” *Synthese*, 53(3): 361–386.

Gelman, Andrew, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. 2013. *Bayesian Data Analysis, Third Edition.* CRC Press. Google-Books-ID: ZXL6AQAAQBAJ.

Giustinelli, Pamela, Charles F Manski, and Francesca Molinari. 2022*a*. “Precise or Imprecise Probabilities? Evidence from Survey Response Related to Late-Onset Dementia.” *Journal of the European Economic Association*, 20(1): 187–221.

Giustinelli, Pamela, Charles F. Manski, and Francesca Molinari. 2022*b*. “Tail and Center Rounding of Probabilistic Expectations in the Health and Retirement Study.” *Journal of Econometrics*, 231(1): 265–281.

Godlonton, Susan, Alister Munthali, and Rebecca Thornton. 2016. "Responding to Risk: Circumcision, Information, and HIV Prevention." *Review of Economics and Statistics*, 98(2): 333–349.

Hainmueller, Jens, and Chad Hazlett. 2014. "Kernel Regularized Least Squares: Reducing Misspecification Bias with a Flexible and Interpretable Machine Learning Approach." *Political Analysis*, 22(2): 143–168.

Hakes, Jahn, and W. Kip Viscusi. 1997. "Mortality Risk Perceptions: A Bayesian Reassessment." *Journal of Risk and Uncertainty*, 15(2): 135–150.

Han, Paul K.J., Sarah C. Kobrin, William M.P. Klein, William W. Davis, Michael Stefanek, and Steven H. Taplin. 2007. "Perceived Ambiguity about Screening Mammography Recommendations: Association with Future Mammography Uptake and Perceptions." *Cancer Epidemiology, Biomarkers & Prevention*, 16(3): 458–466.

Hoel, Jessica B, Hope Michelson, Ben Norton, and Victor Manyong. 2022. "Misattribution and Uncertainty About Beliefs Prevent Learning." Working Paper.

Ilut, Cosmin, and Martin Schneider. 2023. "Chapter 24 - Ambiguity." In *Handbook of Economic Expectations.* , ed. Rüdiger Bachmann, Giorgio Topa and Wilbert van der Klaauw, 749–777. Academic Press.

Imbens, Guido W., and Donald B. Rubin. 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction.* Cambridge:Cambridge University Press.

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica*, 47(2): 263–291.

Kamenica, Emir. 2019. "Bayesian Persuasion and Information Design." *Annual Review of Economics*, 11(1): 249–272.

Kerwin, Jason T. 2025. "Scared Straight or Scared to Death? Fatalism in Response to Disease Risks." *The Economic Journal*, forthcoming.

Knight, Frank Hyneman. 1921. *Risk, Uncertainty, and Profit.* Boston, MA:Houghton Mifflin.

Lybbert, Travis J., Christopher B. Barrett, John G. McPeak, and Winnie K. Luseno. 2007. "Bayesian Herders: Updating of Rainfall Beliefs in Response to External Forecasts." *World Development*, 35(3): 480–497.

Machina, Mark J., and Marciano Siniscalchi. 2014. "Chapter 13 – Ambiguity and Ambiguity Aversion." In *Handbook of the Economics of Risk and Uncertainty.* Vol. 1 of *Handbook of the Economics of Risk and Uncertainty*, , ed. Mark Machina and Kip Viscusi, 729–807. North-Holland.

Maertens, Annemie, Hope Michelson, and Vesall Nourani. 2021. "How Do Farmers Learn from Extension Services? Evidence from Malawi." *American Journal of Agricultural Economics*, 103(2): 569–595.

Malawi National AIDS Commission. 2009. "National HIV/AIDS Prevention Strategy: 2009 to 2013." Malawi National AIDS Commission (NAC), Lilongwe.

Manski, Charles F. 2004. "Measuring Expectations." *Econometrica*, 72(5): 1329–1376.

Nishimura, Kiyohiko G., and Hiroyuki Ozaki. 2004. "Search and Knightian Uncertainty." *Journal of Economic Theory*, 119(2): 299–333.

Paté-Cornell, M. Elisabeth. 1996. "Uncertainties in risk analysis: Six levels of treatment." *Reliability Engineering & System Safety*, 54(2): 95–111.

Pei, Zhuan, Jörn-Steffen Pischke, and Hannes Schwandt. 2019. "Poorly Measured Confounders are More Useful on the Left than on the Right." *Journal of Business & Economic Statistics*, 37(2): 205–216.

Rheinberger, Christoph M., and James K. Hammitt. 2018. "Dinner with Bayes: On the Revision of Risk Beliefs." *Journal of Risk and Uncertainty*, 57(3): 253–280.

Schoemaker, Paul J.H. 1991. "Choices Involving Uncertain Probabilities." *Journal of Economic Behavior & Organization*, 16(3): 295–317.

Sinz, H., L. Zamarian, T. Benke, G. K. Wenning, and M. Delazer. 2008. "Impact of Ambiguity and Risk on Decision Making in Mild Alzheimer's Disease." *Neuropsychologia*, 46(7): 2043–2055.

Sunstein, Cass R., Sebastian Bobadilla-Suarez, Stephanie C. Lazzaro, and Tali Sharot. 2016. "How People Update Beliefs about Climate Change: Good News and Bad News." *Cornell Law Review*, 102(6): 1431–1444.

Tversky, Amos, and Daniel Kahneman. 1992. "Advances in Prospect Theory: Cumulative Representation of Uncertainty." *Journal of Risk and Uncertainty*, 5(4): 297–323.

Viscusi, W. Kip, and Charles J. O'Connor. 1984. "Adaptive Responses to Chemical Labeling: Are Workers Bayesian Decision Makers?" *American Economic Review*, 74(5): 942–956.

Walley, Peter. 1991. *Statistical Reasoning with Imprecise Probabilities. Monographs on Statistics and Applied Probability.* 1st ed ed., London & New York:Chapman and Hall.

Wawer, Maria J., Ronald H. Gray, Nelson K. Sewankambo, David Serwadda, Xianbin Li, Oliver Laeyendecker, Noah Kiwanuka, Godfrey Kigozi, Mohammed Kiddugavu, Thomas Lutalo, Fred Nalugoda, Fred Wabwire-Mangen, Mary P. Meehan, and Thomas C. Quinn. 2005. "Rates of HIV-1 Transmission per Coital Act, by Stage of HIV-1 Infection, in Rakai, Uganda." *Journal of Infectious Diseases*, 191(9): 1403–1409.

Woodford, Michael. 2020. "Modeling Imprecision in Perception, Valuation, and Choice." *Annual Review of Economics*, 12: 579–601.

Yang, Jeffrey. 2023. "On the Decision-Relevance of Subjective Beliefs." *SSRN Electronic Journal.*

Yaniv, Ilan, and Dean P. Foster. 1995. "Graininess of Judgment under Uncertainty: An Accuracy-Informativeness trade-off." *Journal of Experimental Psychology: General*, 124(4): 424–432.

Supplemental Online Appendix, Not Intended for Publication

Appendix A Appendix Tables and Figures

Appendix Table A1

Summary Statistics for Risk Beliefs and Imprecision

(Respondents with Non-zero Imprecision)

				Percentiles										
	N	Mean	SD	Min	1st	5th	10th	25th	50th	75th	90th	95th	99th	Max
HIV Transmission Rate Beliefs														
Annual	298	0.80	0.25	0.07	0.10	0.20	0.40	0.70	0.90	1.00	1.00	1.00	1.00	1.00
Imprecision	298	0.21	0.21	0.01	0.01	0.02	0.03	0.07	0.12	0.29	0.50	0.59	0.98	0.99
Per-act	522	0.67	0.32	0.00	0.02	0.06	0.10	0.50	0.80	0.97	1.00	1.00	1.00	1.00
Imprecision	522	0.20	0.19	0.01	0.01	0.02	0.04	0.06	0.14	0.30	0.50	0.60	0.95	1.00
HIV Prevalence Beliefs														
Attractive People	535	0.56	0.22	0.00	0.00	0.20	0.30	0.40	0.60	0.70	0.80	1.00	1.00	1.00
Imprecision	535	0.24	0.14	0.10	0.10	0.10	0.10	0.10	0.20	0.30	0.40	0.50	0.70	0.90
All People	713	0.52	0.29	0.00	0.02	0.05	0.10	0.25	0.55	0.80	0.90	0.95	0.99	1.00
Imprecision	713	0.16	0.14	0.01	0.01	0.03	0.05	0.05	0.10	0.20	0.35	0.45	0.70	0.89

Notes: Sample is 1,292 people from 70 villages who completed both baseline and endline surveys; for each risk belief, we subset the sample to just those people who report non-zero imprecision levels. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk.

Appendix Table A2

Summary Statistics: Risk Beliefs 100 vs not 100

	Risk < 100 Mean (SD) (1)	Risk = 100 Mean (SD) (2)	Diff. (p-val.) (3)	Obs (4)
Demographics				
Male	0.463	0.415	-0.000	1,284
	(0.499)	(0.493)	()	
Married	0.818	0.815	-0.005	1,282
	(0.386)	(0.389)	(0.853)	
Age	29.129	29.466	0.346	1,284
	(7.980)	(8.534)	(0.444)	
Years of Education	5.313	6.021	0.786***	1,284
	(3.383)	(3.405)	(0.000)	
Household size	4.908	4.969	0.064	1,284
	(1.996)	(2.200)	(0.528)	
Spending in past 30 days	317.634	282.401	-30.187	1,284
	(661.985)	(392.786)	(0.394)	
Assets owned	4.374	4.416	0.084	1,283
	(2.640)	(2.576)	(0.619)	
Raven's score [0-3]	1.437	1.590	0.189***	1,284
	(0.993)	(0.994)	(0.002)	
Numeracy [0-3]	0.697	0.798	0.140**	1,284
	(0.945)	(0.980)	(0.023)	
Risk attitude	1.316	1.249	-0.073***	1,280
	(0.465)	(0.433)	(0.003)	
Christian	0.921	0.920	-0.001	1,284
	(0.270)	(0.271)	(0.953)	
Muslim	0.071	0.071	-0.000	1,284
	(0.257)	(0.257)	(0.980)	
Sexual Activity				
Any Sex in Past Week	0.529	0.522	-0.003	1,284
	(0.500)	(0.500)	(0.902)	
Total Acts in Past Week	1.763	1.685	-0.068	1,284
	(2.497)	(2.404)	(0.679)	
Unprotected Acts in Past Week	1.539	1.514	-0.018	1,284
	(2.366)	(2.348)	(0.914)	
Sex Partners in Past 30 Days	0.863	0.784	-0.075	1,282
	(0.891)	(0.506)	(0.101)	
Condoms Acquired in Past 30 Days	4.842	3.872	-0.869	1,280
	(15.265)	(12.580)	(0.402)	
Years Sexually Active	12.772	13.322	0.530	1,268
	(8.203)	(8.542)	(0.275)	
Lifetime Sex Partners	3.240	3.337	0.208	1,280
	(3.013)	(4.094)	(0.304)	
Any Chance of Having HIV	0.348	0.347	-0.001	1,269
	(0.477)	(0.476)	(0.971)	
Overall Sexual Activity Index	0.021	-0.009	-0.025	1,269
	(0.993)	(1.005)	(0.661)	

Notes: Sample is 1,292 people from 70 villages who completed both baseline and endline surveys. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk.

Appendix Table A3

Predictive Power of Covariates for Imprecision
(Annual Risk Beliefs)

	Outcome: Imprecision in Annual Risk (Standardized)			
	Bivariate Regressions	Multivariable Regression	KRLS	Lasso
Demographics				
Male	0.026	0.035	0.000	
	(0.026)	(0.031)	(0.000)	
Married	0.015	0.026	0.000	
	(0.027)	(0.034)	(0.001)	
Age	-0.020	0.028	-0.000	
	(0.027)	(0.061)	(0.000)	
Years of Education	0.013	0.032	0.000	
	(0.028)	(0.036)	(0.000)	
Household size	0.025	0.036	0.000	
	(0.023)	(0.023)	(0.000)	
Spending in past 30 days	-0.021	-0.017	0.000	
	(0.016)	(0.015)	(0.000)	
Assets owned	-0.037	-0.047	-0.000	
	(0.026)	(0.033)	(0.000)	
Raven's score [0-3]	-0.019	-0.037	-0.000	
	(0.026)	(0.031)	(0.000)	
Numeracy [0-3]	0.012	0.011	0.000	
	(0.032)	(0.043)	(0.000)	
Risk attitude	0.034	0.034	0.000	
	(0.027)	(0.027)	(0.000)	
Christian	-0.036	-0.183	-0.000	
	(0.033)	(0.183)	(0.001)	
Muslim	0.015	-0.155	0.000	
	(0.028)	(0.180)	(0.001)	
Sexual Activity				
Any Sex in Past Week	-0.004	-0.036	-0.000	
	(0.027)	(0.053)	(0.000)	
Total Acts in Past Week	0.000	-0.013	-0.000	
	(0.027)	(0.049)	(0.000)	
Unprotected Acts in Past Week	0.005	0.015	-0.000	
	(0.028)	(0.059)	(0.000)	
Sex Partners in Past 30 Days	0.015	0.017	0.000	
	(0.029)	(0.038)	(0.000)	
Condoms Acquired in Past 30 Days	-0.000	-0.001	0.000	
	(0.019)	(0.021)	(0.000)	
Years Sexually Active	-0.032	-0.045	-0.000	
	(0.028)	(0.061)	(0.000)	
Lifetime Sex Partners	0.008	0.011	0.000	
	(0.027)	(0.026)	(0.000)	
Any Chance of Having HIV	-0.026	-0.018	-0.000	
	(0.026)	(0.026)	(0.000)	
Overall Sexual Activity Index	0.005	0.021	0.000	
	(0.027)	(0.078)	(0.000)	
Observations	1,292	1,237	1,237	1,284
R-squared				
Unadjusted		0.014	0.002	0.000
Adjusted		-0.003	-0.015	0.000

Notes: Sample is 1,292 people from 70 villages who completed both baseline and endline surveys. Differences and p-values in column 3 are adjusted for sampling strata and clustered by village: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A4

Predictive Power of Covariates for Imprecision
(Per-Act Risk Beliefs)

	Outcome: Imprecision in Per-Act Risk (Standardized)			
Variables	Bivariate Regressions	Multivariable Regression	KRLS	Lasso
Demographics				
Male	0.010	-0.050	-0.000	
	(0.030)	(0.040)	(0.000)	
Married	-0.007	0.009	-0.000	
	(0.027)	(0.030)	(0.001)	
Age	0.032	-0.010	0.000	
	(0.025)	(0.066)	(0.000)	
Years of Education	0.016	0.036	0.000	
	(0.029)	(0.037)	(0.000)	
Household size	0.032	0.032	0.000	
	(0.029)	(0.031)	(0.000)	
Spending in past 30 days	-0.011	-0.009	-0.000	
	(0.023)	(0.022)	(0.000)	
Assets owned	-0.004	-0.020	-0.000	
	(0.024)	(0.030)	(0.000)	
Raven's score [0-3]	0.036	0.044	0.000	
	(0.036)	(0.036)	(0.000)	
Numeracy [0-3]	0.008	-0.001	0.000	
	(0.034)	(0.038)	(0.000)	
Risk attitude	0.067*	0.077**	0.000	0.073**
	(0.035)	(0.034)	(0.000)	(0.034)
Christian	0.020	0.035	0.000	
	(0.021)	(0.051)	(0.001)	
Muslim	-0.021	0.022	-0.000	
	(0.021)	(0.052)	(0.001)	
Sexual Activity				
Any Sex in Past Week	0.031	0.064	0.000	
	(0.028)	(0.044)	(0.000)	
Total Acts in Past Week	0.035	-0.020	0.000	
	(0.029)	(0.048)	(0.000)	
Unprotected Acts in Past Week	0.039	0.097	0.000	
	(0.029)	(0.064)	(0.000)	
Sex Partners in Past 30 Days	-0.016	-0.005	-0.000	
	(0.021)	(0.029)	(0.000)	
Condoms Acquired in Past 30 Days	-0.002	-0.003	-0.000	
	(0.036)	(0.038)	(0.000)	
Years Sexually Active	0.047	0.023	0.000*	
	(0.029)	(0.074)	(0.000)	
Lifetime Sex Partners	0.142**	0.149*	0.000*	0.146**
	(0.067)	(0.077)	(0.000)	(0.067)
Any Chance of Having HIV	0.050*	0.029	0.000	
	(0.028)	(0.031)	(0.001)	
Overall Sexual Activity Index	0.013	-0.109	0.000	
	(0.029)	(0.073)	(0.000)	
Observations	1,282	1,239	1,239	1,280
R-squared				
Unadjusted		0.035	0.002	0.026
Adjusted		0.019	-0.015	0.024

Notes: Sample is 1,292 people from 70 villages who completed both baseline and endline surveys. Differences and p-values in column 3 are adjusted for sampling strata and clustered by village: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A5

Predictive Power of Covariates for Non-zero Imprecision
(Annual Risk Beliefs)

	Outcome: Imprecision in Annual Risk						
	Bivariate Regressions	Multivariable Regression	Probit	Logit	Tobit	KRLS	Lasso
Demographics							
Male	0.033**	0.040**	0.081***	0.458***	0.079**	0.004**	0.037***
	(0.013)	(0.015)	(0.030)	(0.177)	(0.034)	(0.002)	(0.013)
Married	0.007	0.004	0.013	0.075	0.026	-0.000	
	(0.012)	(0.014)	(0.037)	(0.213)	(0.043)	(0.002)	
Age	0.008	0.019	0.002	0.012	0.002	0.000	
	(0.011)	(0.026)	(0.003)	(0.017)	(0.004)	(0.000)	
Years of Education	-0.011	-0.015	-0.005	-0.026	-0.001	-0.000	
	(0.011)	(0.014)	(0.004)	(0.024)	(0.005)	(0.000)	
Household size	0.022*	0.025**	0.011**	0.062**	0.012*	0.000*	0.023*
	(0.011)	(0.012)	(0.005)	(0.029)	(0.007)	(0.000)	(0.011)
Spending in past 30 days	-0.001	-0.001	-0.000	-0.000	-0.000	0.000	
	(0.008)	(0.008)	(0.000)	(0.000)	(0.000)	(0.000)	
Assets owned	-0.005	-0.012	-0.004	-0.025	-0.008	-0.000	
	(0.014)	(0.016)	(0.006)	(0.035)	(0.006)	(0.000)	
Raven's score [0-3]	0.004	0.003	0.003	0.017	-0.007	0.000	
	(0.012)	(0.014)	(0.014)	(0.078)	(0.017)	(0.001)	
Numeracy [0-3]	0.010	0.004	0.004	0.027	0.006	0.001	
	(0.013)	(0.017)	(0.017)	(0.098)	(0.018)	(0.001)	
Risk attitude	0.026**	0.028**	0.064**	0.361**	0.066**	0.000	0.031**
	(0.012)	(0.012)	(0.027)	(0.149)	-0.033	(0.001)	(0.012)
Christian	-0.019*	-0.006	-0.019	-0.096	-0.151	-0.003	-0.019*
	(0.011)	(0.038)	(0.125)	(0.695)	(0.141)	(0.003)	(0.011)
Muslim	0.019	0.014	0.050	0.302	-0.086	0.004	
	(0.012)	(0.041)	(0.138)	(0.763)	(0.149)	(0.004)	
Sexual Activity							
Any Sex in Past Week	0.014	-0.016	-0.032	-0.178	-0.041	0.000	
	(0.012)	(0.023)	(0.045)	(0.253)	(0.054)	(0.002)	
Total Acts in Past Week	0.015	0.021	0.008	0.045	0.004	0.000	
	(0.014)	(0.042)	(0.015)	(0.083)	(0.017)	(0.000)	
Unprotected Acts in Past Week	0.013	-0.018	-0.007	-0.041	-0.002	0.000	
	(0.014)	(0.040)	(0.015)	(0.083)	(0.018)	(0.000)	
Sex Partners in Past 30 Days	0.012	0.001	0.005	0.013	0.011	0.001*	
	(0.011)	(0.013)	(0.025)	(0.127)	(0.033)	(0.001)	
Condoms Acquired in Past 30 Days	0.014	0.009	0.001	0.003	0.001	0.000	
	(0.015)	(0.017)	(0.001)	(0.005)	(0.001)	(0.000)	
Years Sexually Active	0.006	-0.016	-0.002	-0.010	-0.003	0.000	
	(0.012)	(0.027)	(0.003)	(0.018)	(0.004)	(0.000)	
Lifetime Sex Partners	0.013	0.005	0.002	0.010	0.002	0.000***	
	(0.009)	(0.009)	(0.003)	(0.016)	(0.004)	(0.000)	
Any Chance of Having HIV	0.012	0.010	0.021	0.120	0.006	0.002	
	(0.014)	(0.013)	(0.028)	(0.156)	(0.031)	(0.002)	
Overall Sexual Activity Index	0.015	0.022	0.019	0.122	0.017	0.000	
	(0.013)	(0.034)	(0.033)	(0.184)	(0.044)	(0.000)	
Observations	1,292	1,237	1,237	1,237	1,237	1,237	1,280
R-squared							
Unadjusted		0.024				0.006	0.016
Adjusted		0.007				-0.011	0.013
Pseudo			0.022	0.022	0.021		
Adjusted Pseudo			-0.011	-0.011	-0.020		
Control-group Mean	0.232	0.229	0.229	0.229	0.229	0.229	0.229
Control-group SD	0.422	0.420	0.420	0.420	0.420	0.420	0.421

Notes: Sample is 1,292 people from 70 villages who completed both baseline and endline surveys. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk.

Appendix Table A6

Predictive Power of Covariates for Non-zero Imprecision
(Per-Act Risk Beliefs)

	Outcome: Imprecision in Per-Act Risk					
Variables	Bivariate Regressions	Multivariable Regression	Probit	Logit	KRLS	Lasso
Demographics						
Male	0.040**	0.022	0.045	0.044	0.004	0.030*
	(0.015)	(0.018)	(0.037)	(0.037)	(0.002)	(0.016)
Married	0.012	0.011	0.027	0.028	0.001	
	(0.014)	(0.016)	(0.043)	(0.043)	(0.003)	
Age	0.012	0.020	0.002	0.002	0.000*	
	(0.014)	(0.031)	(0.004)	(0.004)	(0.000)	
Years of Education	-0.011	-0.029	-0.009	-0.009	-0.000	
	(0.014)	(0.018)	(0.005)	(0.005)	(0.000)	
Household size	0.011	0.009	0.004	0.004	0.000	
	(0.013)	(0.013)	(0.006)	(0.006)	(0.000)	
Spending in past 30 days	-0.007	-0.011	-0.000	-0.000	-0.000	
	(0.009)	(0.008)	(0.000)	(0.000)	(0.000)	
Assets owned	0.020	0.023	0.009	0.009	0.000	
	(0.015)	(0.016)	(0.006)	(0.006)	(0.000)	
Raven's score [0-3]	0.023*	0.021	0.022	0.022	0.001	
	(0.013)	(0.015)	(0.015)	(0.015)	(0.001)	
Numeracy [0-3]	0.021*	0.012	0.012	0.013	0.001	
	(0.012)	(0.015)	(0.016)	(0.017)	(0.001)	
Risk attitude	0.025	0.035**	0.081**	0.081**	0.001	0.031*
	(0.016)	(0.016)	(0.038)	(0.038)	(0.002)	(0.016)
Christian	-0.025*	-0.049	-0.184	-0.179	-0.003	-0.026**
	(0.013)	(0.037)	(0.147)	(0.147)	(0.004)	(0.012)
Muslim	0.019	-0.025	-0.100	-0.095	0.003	
	(0.013)	(0.037)	(0.148)	(0.148)	(0.004)	
Sexual Activity						
Any Sex in Past Week	0.023	0.004	0.008	0.007	0.002	
	(0.015)	(0.025)	(0.050)	(0.050)	(0.002)	
Total Acts in Past Week	0.022	0.008	0.003	0.003	0.000	
	(0.014)	(0.045)	(0.018)	(0.018)	(0.000)	
Unprotected Acts in Past Week	0.023*	0.013	0.005	0.005	0.000	
	(0.013)	(0.043)	(0.018)	(0.018)	(0.000)	
Sex Partners in Past 30 Days	0.005	-0.006	-0.014	-0.014	0.001	
	(0.012)	(0.011)	(0.028)	(0.028)	(0.001)	
Condoms Acquired in Past 30 Days	-0.001	-0.005	-0.000	-0.000	-0.000	
	(0.014)	(0.015)	(0.001)	(0.001)	(0.000)	
Years Sexually Active	0.013	-0.026	-0.003	-0.003	0.000	
	(0.014)	(0.032)	(0.004)	(0.004)	(0.000)	
Lifetime Sex Partners	0.036***	0.027**	0.011**	0.011**	0.000***	0.031***
	(0.010)	(0.010)	(0.005)	(0.005)	(0.000)	(0.010)
Any Chance of Having HIV	0.023*	0.019	0.040	0.041	0.003	
	(0.014)	(0.014)	(0.030)	(0.030)	(0.002)	
Overall Sexual Activity Index	0.020	-0.008	-0.007	-0.007	0.001	
	(0.015)	(0.037)	(0.038)	(0.038)	(0.000)	
Observations	1,292	1,239	1,239	1,239	1,239	1,280
R-squared						
Unadjusted		0.029			0.006	0.020
Adjusted		0.013			-0.011	0.017
Pseudo			0.022	0.022		
Adjusted Pseudo			-0.004	-0.004		
Control-group Mean	0.406	0.403	0.403	0.403	0.403	0.402
Control-group SD	0.491	0.491	0.491	0.491	0.491	0.491

Notes: Sample is 1,292 people from 70 villages who completed both baseline and endline surveys. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk.

Appendix Table A7

Predictive Power of Covariates for Imprecision
(Tobit Analysis)

	Imprecision in Annual Risk	*Imprecision in Per-act Risk*
Demographics		
Male	0.079**	-0.001
	(0.034)	(0.024)
Married	0.026	0.014
	(0.043)	(0.030)
Age	0.002	0.001
	(0.004)	(0.003)
Years of Education	-0.001	-0.001
	(0.005)	(0.004)
Household size	0.012*	0.005
	(0.007)	(0.005)
Spending in past 30 days	-0.000	-0.000
	(0.000)	(0.000)
Assets owned	-0.008	0.002
	(0.006)	(0.004)
Raven's score [0-3]	-0.007	0.017
	(0.017)	(0.012)
Numeracy [0-3]	0.006	0.003
	(0.018)	(0.012)
Risk attitude	0.066**	0.065***
	-0.033	(0.023)
Christian	-0.151	-0.025
	(0.141)	(0.100)
Muslim	-0.086	-0.003
	(0.149)	(0.105)
Sexual Activity		
Any Sex in Past Week	-0.041	0.029
	(0.054)	(0.038)
Total Acts in Past Week	0.004	-0.000
	(0.017)	(0.012)
Unprotected Acts in Past Week	-0.002	0.009
	(0.018)	(0.013)
Sex Partners in Past 30 Days	0.011	-0.013
	(0.033)	(0.029)
Condoms Acquired in Past 30 D	0.001	-0.000
	(0.001)	(0.001)
Years Sexually Active	-0.003	-0.000
	(0.004)	(0.003)
Lifetime Sex Partners	0.002	0.010***
	(0.004)	(0.003)
Any Chance of Having HIV	0.006	0.028
	(0.031)	(0.021)
Overall Sexual Activity Index	0.017	-0.021
	(0.044)	(0.032)
Observations	1,237	1,239
R-squared		
Pseudo	0.021	0.035
Adjusted Pseudo	-0.024	-0.007
Control-group Mean	0.229	0.228
Control-group SD	0.420	0.420

Notes: Sample is 1,292 people from 70 villages who completed both baseline and endline surveys. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk.

Appendix Table A8

Bins for Annual and Per-Act Risk Beliefs

Bin	N	Min	Max
	Annual Risk Beliefs		
1	141	0.04	0.60
2	124	0.65	0.80
3	1,019	0.85	1.00
	Per-Act Risk Beliefs		
1	129	0.00	0.22
2	173	0.25	0.50
3	105	0.55	0.70
4	110	0.75	0.89
5	137	0.90	0.98
6	635	0.99	1.00

Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey. Both variables are measured at baseline. Annual risk beliefs are the perceived chance of contracting HIV from regular unprotected sex with an HIV-positive sexual partner over a period of one year. Per-act risk beliefs are the perceived chance of contracting HIV from a single unprotected sex act with an HIV-positive sexual partner.

Appendix Table A9

Treatment Effect Heterogeneity by Imprecision and Brackets of Beliefs

	Outcome: Endline Annual HIV Transmission Risk			
	(1)	(2)	(3)	(4)
Treatment ×				
Bottom Bin of Baseline Beliefs	-0.173***	-0.176***	-0.174***	-0.168***
	(0.057)	(0.058)	(0.057)	(0.058)
Middle Bin of Baseline Beliefs	-0.425***	-0.394***	-0.417***	-0.385***
	(0.051)	(0.057)	(0.054)	(0.061)
Top Bin of Baseline Beliefs	-0.391***	-0.392***	-0.378***	-0.381***
	(0.018)	(0.018)	(0.017)	(0.016)
Imprecision		-0.020		0.002
		(0.081)		(0.079)
Treatment × Imprecision		0.075		-0.031
		(0.230)		(0.244)
Treatment × Imprecision x				
Bottom Bin of Baseline Beliefs		0.000		
		(0.000)		
Middle Bin of Baseline Beliefs		-0.599*		-0.536
		(0.335)		(0.367)
Top Bin of Baseline Beliefs		-0.353		-0.169
		(0.219)		(0.254)
T Interacted w/ Other Baseline Covariates	No	No	Yes	Yes
Observations	1,276	1,276	1,230	1,230
Adjusted R-squared	0.334	0.340	0.389	0.389
Control-group Means				
Overall	0.906	0.906	0.907	0.907
Bottom Bin	0.643	0.643	0.640	0.640
Middle Bin	0.847	0.847	0.842	0.842
Top Bin	0.937	0.937	0.938	0.938
Control-group SD	0.196	0.196	0.195	0.195

Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk. Other baseline covariates include all variables in Table 1. All regressions control for stratification cell fixed effects. Heteroskedasticity-robust standard errors, clustered by village, in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A10

Treatment Effect Heterogeneity by Imprecision

(Dropping Highest and Lowest Risk Beliefs)

	Outcome: Endline Annual HIV Transmission Risk				*Outcome: Endline Per-Act HIV Transmission Risk*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment Group	-0.345***	-0.328***	-0.453***	-0.409***	-0.343***	-0.340***	-0.405***	-0.368***
	(0.030)	(0.033)	(0.034)	(0.042)	(0.022)	(0.023)	(0.027)	(0.027)
Imprecision		0.185	0.211*	0.247*		-0.076	-0.035	-0.097
		(0.112)	(0.107)	(0.131)		(0.123)	(0.116)	(0.116)
Baseline Beliefs	0.184***	0.189***	0.558***	0.504***	0.191***	0.193***	0.404***	0.353***
	(0.065)	(0.064)	(0.090)	(0.093)	(0.045)	(0.045)	(0.064)	(0.058)
Treatment × Imprecision		-0.342*	-0.349**	-0.423**		-0.028	-0.056	-0.096
		(0.175)	(0.167)	(0.195)		(0.154)	(0.148)	(0.159)
Treatment × Baseline Beliefs			-0.578***	-0.469***			-0.356***	-0.260***
			(0.109)	(0.128)			(0.085)	(0.082)
T Interacted w/ Other Baseline Covariates	No	No	No	Yes	No	No	No	Yes
Observations	378	378	378	366	675	673	673	650
Adjusted R-squared	0.257	0.258	0.292	0.308	0.268	0.268	0.287	0.305
Control-group Mean	0.800	0.800	0.800	0.797	0.645	0.645	0.645	0.644
Control-group SD	0.261	0.261	0.261	0.263	0.320	0.321	0.321	0.322

Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey; for this analysis we drop respondents with risk beliefs equal to 0% or 100%. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk. Other baseline covariates include all variables in Table 1. All regressions control for stratification cell fixed effects. Heteroskedasticity-robust standard errors, clustered by village, in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A11

Treatment Effect Heterogeneity by Imprecision
(Adjusting for Heterogeneity by Brackets of Beliefs)
Dropping Highest and Lowest Risk Beliefs

	Outcome: Endline Annual HIV Transmission			
	(1)	(2)	(3)	(4)
Treatment ×				
Bottom Bin of Baseline Beliefs	-0.176***	-0.160***	-0.180***	-0.171***
	(0.057)	(0.058)	(0.059)	(0.061)
Middle Bin of Baseline Beliefs	-0.426***	-0.403***	-0.427***	-0.406***
	(0.052)	(0.056)	(0.053)	(0.056)
Top Bin of Baseline Beliefs	-0.456***	-0.448***	-0.410***	-0.401***
	(0.039)	(0.040)	(0.050)	(0.051)
Imprecision		0.199		0.229
		(0.122)		(0.146)
Treatment × Imprecision		-0.340*		-0.426**
		(0.179)		(0.199)
T Interacted w/ Other Baseline Covariates	No	No	Yes	Yes
Observations	378	378	366	366
Adjusted R-squared	0.276	0.277	0.293	0.282
Control-group Means				
Overall	0.800	0.800	0.797	0.797
Bottom Bin	0.643	0.643	0.640	0.640
Middle Bin	0.847	0.847	0.842	0.842
Top Bin	0.884	0.884	0.885	0.885
Control-group SD	0.261	0.261	0.263	0.263

Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey; for this analysis we drop respondents with risk beliefs equal to 0% or 100%. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk. Other baseline covariates include all variables in Table 1. All regressions control for stratification cell fixed effects. Heteroskedasticity-robust standard errors, clustered by village, in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A12

Treatment Effect Heterogeneity by Imprecision
(Dropping observations with Zero Imprecision)

	Outcome: Endline Annual HIV Transmission Risk				*Outcome: Endline Per-Act HIV Transmission Risk*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment Group	-0.411***	-0.382***	-0.438***	-0.383***	-0.416***	-0.432***	-0.478***	-0.435***
	(0.033)	(0.042)	(0.041)	(0.049)	(0.026)	(0.034)	(0.035)	(0.038)
Imprecision		0.062	0.083	0.125		-0.119	-0.090	-0.075
		(0.086)	(0.089)	(0.109)		(0.121)	(0.124)	(0.123)
Baseline Beliefs	0.137*	0.146*	0.450***	0.375***	0.191***	0.189***	0.412***	0.383***
	(0.081)	(0.082)	(0.108)	(0.112)	(0.052)	(0.052)	(0.073)	(0.078)
Treatment × Imprecision		-0.175	-0.162	-0.311*		0.151	0.175	0.071
		(0.138)	(0.136)	(0.175)		(0.151)	(0.151)	(0.155)
Treatment × Baseline Beliefs			-0.515***	-0.346**			-0.396***	-0.318***
			(0.145)	(0.149)			(0.093)	(0.106)
T Interacted w/ Other Baseline Covariates	No	No	No	Yes	No	No	No	Yes
Observations	294	294	294	283	522	520	520	509
Adjusted R-squared	0.340	0.339	0.364	0.342	0.363	0.362	0.386	0.390
Control-group Mean	0.845	0.845	0.845	0.844	0.693	0.693	0.693	0.693
Control-group SD	0.239	0.239	0.239	0.242	0.325	0.326	0.326	0.325

Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey; for this analysis we drop respondents with zero imprecision. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk. Other baseline covariates include all variables in Table 1. All regressions control for stratification cell fixed effects. Heteroskedasticity-robust standard errors, clustered by village, in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A13

Treatment Effect Heterogeneity by Imprecision
(Adjusting for Heterogeneity by Brackets of Beliefs)
Dropping observations with Zero Imprecision

	Outcome: Endline Annual HIV Transmission			
	(1)	(2)	(3)	(4)
Treatment ×				
Bottom Bin of Baseline Beliefs	-0.168**	-0.139	-0.217**	-0.172*
	(0.083)	(0.085)	(0.087)	(0.092)
Middle Bin of Baseline Beliefs	-0.488***	-0.455***	-0.478***	-0.418***
	(0.066)	(0.074)	(0.084)	(0.093)
Top Bin of Baseline Beliefs	-0.475***	-0.446***	-0.440***	-0.386***
	(0.038)	(0.045)	(0.049)	(0.055)
Imprecision		0.095		0.143
		(0.095)		(0.116)
Treatment × Imprecision		-0.191		-0.330*
		(0.140)		(0.178)
T Interacted w/ Other Baseline Covariates	No	No	Yes	Yes
Observations	294	294	283	283
Adjusted R-squared	0.369	0.367	0.343	0.348
Control-group Means				
Overall	0.845	0.845	0.844	0.844
Bottom Bin	0.659	0.659	0.653	0.653
Middle Bin	0.813	0.813	0.813	0.813
Top Bin	0.903	0.903	0.902	0.902
Control-group SD	0.239	0.239	0.242	0.242

Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey; for this analysis we drop respondents with zero imprecision. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk. Other baseline covariates include all variables in Table 1. All regressions control for stratification cell fixed effects. Heteroskedasticity-robust standard errors, clustered by village, in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A14

Treatment Effect Heterogeneity by Imprecision
(Half-Width Instead of Range)

	Outcome: Endline Annual HIV Transmission Risk				*Outcome: Endline Per-Act HIV Transmission Risk*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment Group	-0.371***	-0.369***	-0.367***	-0.358***	-0.384***	-0.381***	-0.379***	-0.374***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.019)	(0.019)	(0.019)	(0.018)
Imprecision		-0.092	-0.019	0.005		-0.155*	-0.078	-0.094
		(0.082)	(0.086)	(0.082)		(0.090)	(0.096)	(0.086)
Baseline Beliefs	0.274***	0.258***	0.498***	0.477***	0.272***	0.262***	0.447***	0.429***
	(0.048)	(0.048)	(0.060)	(0.057)	(0.033)	(0.033)	(0.047)	(0.052)
Treatment × Imprecision		-0.179	-0.275**	-0.229*		0.015	-0.082	-0.072
		(0.132)	(0.133)	(0.125)		(0.115)	(0.120)	(0.121)
Treatment × Baseline Beliefs			-0.368***	-0.340***			-0.301***	-0.282***
			(0.081)	(0.079)			(0.060)	(0.067)
T Interacted w/ Other Baseline Covariates	No	No	No	Yes	No	No	No	Yes
Observations	1,276	1,276	1,276	1,230	1,281	1,281	1,281	1,234
Adjusted R-squared	0.328	0.333	0.342	0.393	0.315	0.317	0.329	0.355
Control-group Mean	0.906	0.906	0.906	0.907	0.743	0.743	0.743	0.742
Control-group SD	0.196	0.196	0.196	0.195	0.317	0.317	0.317	0.317

Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey. Imprecision is measured as the "half-width" of the range that the respondent places around their point estimate of the risk. This is the maximum of two numbers—how much lower and how much higher the risk can go. Other baseline covariates include all variables in Table 1. All regressions control for stratification cell fixed effects. Heteroskedasticity-robust standard errors, clustered by village, in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A15

Treatment Effect Heterogeneity by Imprecision

(Truncated Half-Width Instead of Range)

	Outcome: Endline Annual HIV Transmission Risk				*Outcome: Endline Per-Act HIV Transmission Risk*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment Group	-0.371***	-0.369***	-0.367***	-0.358***	-0.384***	-0.382***	-0.379***	-0.375***
	(0.016)	(0.016)	(0.017)	(0.016)	(0.019)	(0.019)	(0.019)	(0.018)
Imprecision		-0.153	-0.040	-0.002		-0.158	-0.057	-0.082
		(0.111)	(0.124)	(0.118)		(0.099)	(0.100)	(0.093)
Baseline Beliefs	0.274***	0.249***	0.495***	0.477***	0.272***	0.260***	0.449***	0.430***
	(0.048)	(0.048)	(0.062)	(0.059)	(0.033)	(0.034)	(0.047)	(0.053)
Treatment × Imprecision		-0.193	-0.355*	-0.318*		-0.017	-0.157	-0.134
		(0.173)	(0.179)	(0.170)		(0.134)	(0.136)	(0.140)
Treatment × Baseline Beliefs			-0.377***	-0.350***			-0.308***	-0.288***
			(0.082)	(0.081)			(0.060)	(0.068)
T Interacted w/ Other Baseline Covariates	No	No	No	Yes	No	No	No	Yes
Observations	1,276	1,276	1,276	1,230	1,281	1,281	1,281	1,234
Adjusted R-squared	0.328	0.333	0.342	0.394	0.315	0.317	0.329	0.355
Control-group Mean	0.906	0.906	0.906	0.742	0.743	0.743	0.743	0.742
Control-group SD	0.196	0.196	0.196	0.317	0.317	0.317	0.317	0.317

Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey. Imprecision is measured as the "truncated half-width" of the range that the respondent places around their point estimate of the risk. This is the maximum of two numbers—how much lower and how much higher the risk can go—but top-coded at a value of 50 percent. Other baseline covariates include all variables in Table 1. All regressions control for stratification cell fixed effects. Heteroskedasticity-robust standard errors, clustered by village, in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A16

Treatment Effect Heterogeneity by Imprecision

(Defining Imprecision as a Binary Indicator)

	Outcome: Endline Annual HIV Transmission Risk				*Outcome: Endline Per-Act HIV Transmission Risk*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment Group	-0.371***	-0.372***	-0.370***	-0.360***	-0.384***	-0.386***	-0.383***	-0.379***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.019)	(0.018)	(0.019)	(0.018)
Any Imprecise Probability		-0.043*	-0.013	-0.024		-0.039	-0.006	-0.026
		(0.022)	(0.022)	(0.021)		(0.024)	(0.025)	(0.026)
Baseline Beliefs	0.274***	0.230***	0.491***	0.460***	0.272***	0.242***	0.451***	0.424***
	(0.048)	(0.049)	(0.060)	(0.057)	(0.033)	(0.036)	(0.048)	(0.055)
Treatment × Imprecision		-0.047	-0.101**	-0.064		-0.038	-0.101***	-0.073*
		(0.041)	(0.041)	(0.039)		(0.033)	(0.036)	(0.038)
Treatment × Baseline Beliefs			-0.405***	-0.358***			-0.349***	-0.320***
			(0.082)	(0.079)			(0.063)	(0.071)
T Interacted w/ Other Baseline Covariates	No	No	No	Yes	No	No	No	Yes
Observations	1,276	1,276	1,276	1,230	1,281	1,279	1,279	1,232
Adjusted R-squared	0.328	0.333	0.344	0.395	0.315	0.320	0.334	0.359
Control-group Mean	0.906	0.906	0.906	0.907	0.743	0.743	0.743	0.742
Control-group SD	0.196	0.196	0.196	0.195	0.317	0.317	0.317	0.317

Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey. Imprecision is measured as an indicator that is equal to 1 if respondents have any imprecision and 0 otherwise. Other baseline covariates include all variables in Table 1. All regressions control for stratification cell fixed effects. Heteroskedasticity-robust standard errors, clustered by village, in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A17

Sexual Activity and Baseline Imprecision

	Outcome: Baseline Sexual Activity (IHS Sex Acts in Past Week)					
	(1)	(2)	(3)	(4)	(5)	(6)
Baseline Imprecision in Annual Risk	-0.042	0.161	0.172			
	(0.387)	(0.392)	(0.406)			
Baseline Imprecision in Per-Act Risk				0.094	0.031	0.031
				(0.300)	(0.242)	(0.245)
Baseline Beliefs		-0.238	0.232		-0.240*	0.465
		(0.208)	(0.787)		(0.119)	(0.863)
Treatment x Brackets of Baseline Beliefs	No	No	Yes	No	No	Yes
Other Baseline Covariates	No	Yes	Yes	No	Yes	Yes
Observations	643	639	639	643	638	638
Adjusted R-squared	-0.003	0.140	0.138	-0.002	0.144	0.139
Control-group Mean	0.933	0.933	0.933	0.933	0.934	0.934
Control-group SD	0.977	0.978	0.978	0.977	0.978	0.978

Notes: Sample is 1,292 people from 70 villages who completed both baseline and endline surveys. Baseline covariates in this table include only demographic characteristics. Differences and p-values in column 3 are adjusted for sampling strata and clustered by village: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A18

Baseline and Endline Imprecision

	Outcome: Endline Imprecision in Annual Risk		*Outcome: Endline Imprecision in Per-Act Risk*	
	(1)	(2)	(3)	(4)
Baseline Imprecision in Annual Risk	0.225***	0.194***		
	(0.069)	(0.071)		
Baseline Imprecision in Per-Act Risk			0.142**	0.148**
			(0.055)	(0.063)
Baseline Beliefs	No	Yes	No	Yes
Other Baseline Covariates	No	Yes	No	Yes
Observations	636	623	635	622
Adjusted R-squared	0.052	0.069	0.011	0.014
Control-group Mean	0.041	0.040	0.074	0.073
Control-group SD	0.105	0.105	0.144	0.142

Notes: Sample is 1,292 people from 70 villages who completed both baseline and endline surveys. Differences and p-values in column 3 are adjusted for sampling strata and clustered by village: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A19

Effects on Endline Sexual Activity

	Outcome: Endline Sexual Activity (IHS Sex Acts in Past Week)						*Outcome: Endline Sexual Activity (IHS Sex Acts in Past Week)*					
	Annual Risk Beliefs and Imprecision in Annual Risk Beliefs						*Per-act Risk Beliefs and Imprecision in Per-Act Risk Beliefs*					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treatment Group	0.104**	0.101**	0.100**	0.092*	0.102**	0.096*	0.125***	0.124***	0.128***	0.116**	0.129***	0.120**
	(0.046)	(0.045)	(0.045)	(0.049)	(0.045)	(0.049)	(0.046)	(0.046)	(0.045)	(0.049)	(0.046)	(0.050)
Imprecision		0.302	0.353	0.408	0.334	0.395		0.009	0.128	0.254	0.143	0.265
		(0.331)	(0.342)	(0.340)	(0.336)	(0.335)		(0.278)	(0.301)	(0.316)	(0.297)	(0.311)
Baseline Beliefs	-0.013	0.020	0.073	0.123	0.120	0.178	0.175*	0.174*	0.463***	0.452***	0.490***	0.484***
	(0.130)	(0.135)	(0.232)	(0.227)	(0.232)	(0.224)	(0.089)	(0.091)	(0.124)	(0.124)	(0.123)	(0.123)
Treatment × Imprecision		-0.067	-0.133	-0.300	-0.111	-0.291		-0.059	-0.221	-0.370	-0.234	-0.380
		(0.369)	(0.387)	(0.401)	(0.380)	(0.397)		(0.330)	(0.352)	(0.368)	(0.349)	(0.365)
Treatment × Baseline Beliefs			-0.067	-0.057	-0.156	-0.160			-0.473***	-0.438**	-0.512***	-0.484***
			(0.290)	(0.295)	(0.272)	(0.274)			(0.169)	(0.181)	(0.166)	(0.179)
Treatment × Baseline Prevalence			-0.375**	-0.363**					-0.284	-0.267		
			(0.185)	(0.176)					(0.190)	(0.181)		
Treatment × Baseline Beliefs × Baseline Prevalence			0.211	0.430					0.023	0.211		
			(0.789)	(0.818)					(0.566)	(0.532)		
Baseline Total Sexual Acts in Past Week	0.520***	0.519***	0.518***	0.596***	0.519***	0.593***	0.525***	0.525***	0.528***	0.622***	0.530***	0.621***
	(0.026)	(0.026)	(0.026)	(0.197)	(0.026)	(0.196)	(0.025)	(0.025)	(0.025)	(0.193)	(0.026)	(0.193)
T Interacted w/ Other Baseline Covariates	No	No	No	Yes	No	Yes	No	No	No	Yes	No	Yes
Observations	1,273	1,273	1,273	1,229	1,273	1,229	1,275	1,274	1,274	1,230	1,274	1,230
Adjusted R-squared	0.278	0.278	0.279	0.294	0.278	0.292	0.282	0.281	0.286	0.301	0.286	0.301
Control-group Mean	0.864	0.864	0.864	0.870	0.864	0.870	0.864	0.866	0.866	0.872	0.866	0.872
Control-group SD	0.980	0.980	0.980	0.980	0.980	0.980	0.980	0.980	0.980	0.980	0.980	0.980

Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk. Other baseline covariates include all variables in Table 1. All regressions control for stratification cell fixed effects. Heteroskedasticity-robust standard errors, clustered by village, in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Figure A1

Survey Question for Unprotected Annual HIV Transmission Risks
Including Followup Questions about 50% Chances and Imprecision

H1a. If 100 women, who do **not** have HIV, each have an HIV-positive sex partner for **one year,** and do **not** use condoms when having sex, how many of the women do you think will have HIV at the end of the year?

Number: | # | # | # |

H1b. **If answer to H1a is 50** Do you really think that 50 of the women 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? | # | # | # |

H1c. **If answer to H1a is <u>not</u> 100** Could it be more? If so, what is the most women out of 100 that you think could possibly get the virus?

Number: | # | # | # |

Interviewer check: Answer must be at least as large as H1a. The same answer is okay.

☐ 997. No, it could not be more.

H1d. **If answer to H1a is <u>not</u> 0** Could it be less? If so, what is the fewest women out of 100 that you think could possibly get the virus?

Number: | # | # | # |

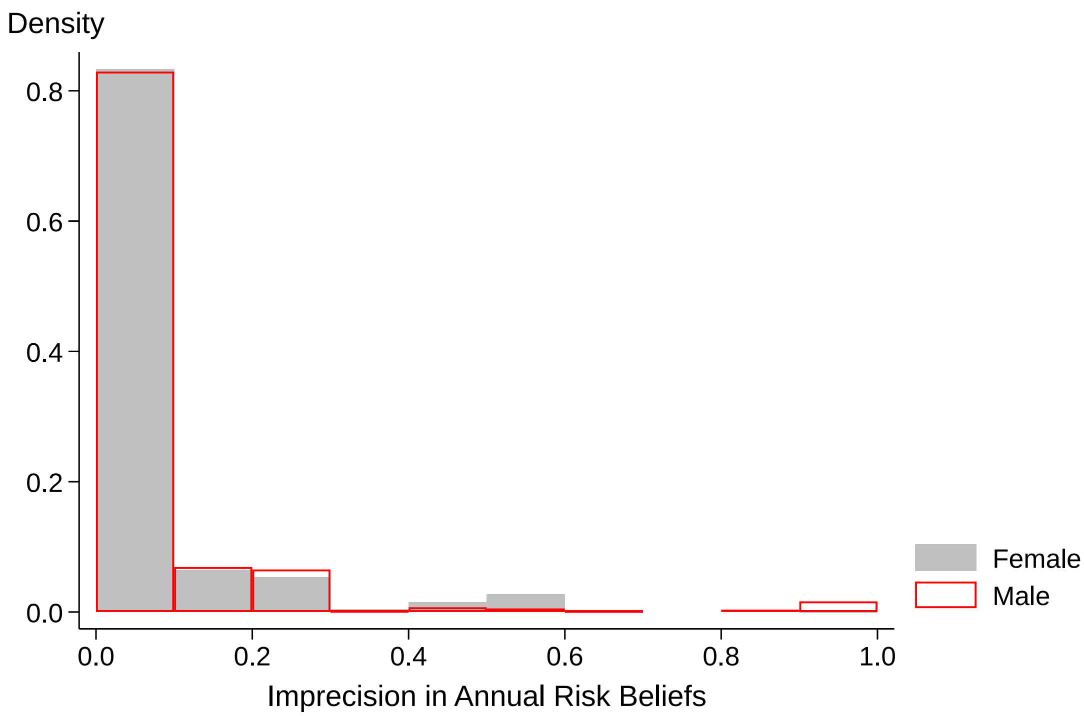
Interviewer check: Answer must be no larger than H1a. The same answer is okay.

☐ 997. No, it could not be less.

Notes: This is the female-specific version of the question; male questions inverted the gender. Questions with parallel phrasing and followups were also used to collected per-act risk beliefs, and beliefs about condom-protected transmission risks.

Appendix Figure A2

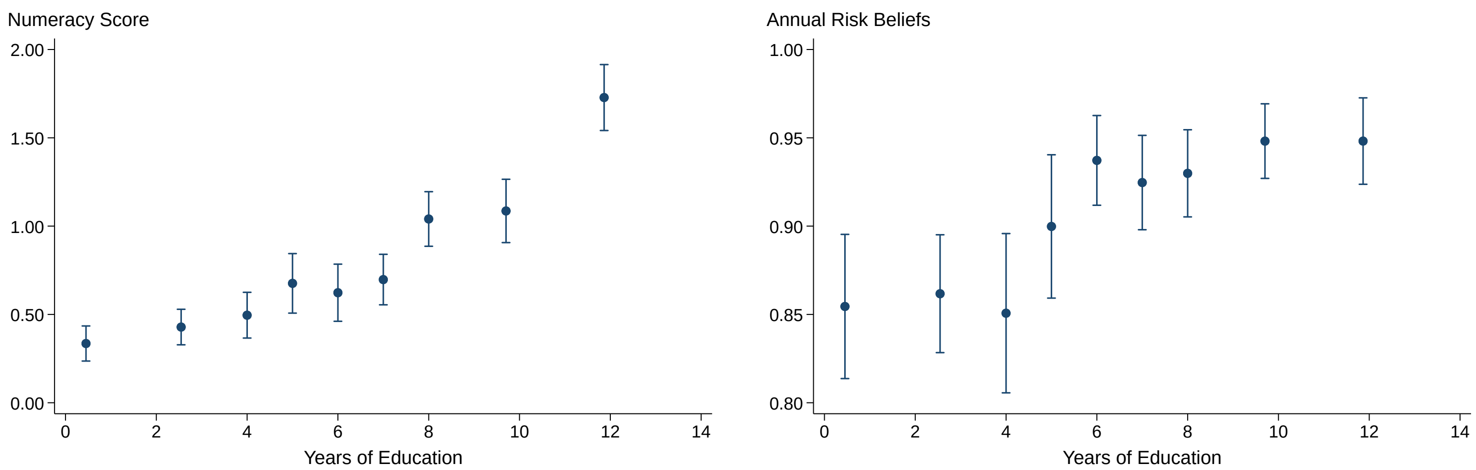
Histogram of Imprecision by Gender
(Annual Risk Beliefs)



Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey. Both variables are measured at baseline. Annual risk beliefs are the perceived chance of contracting HIV from regular unprotected sex with an HIV-positive sexual partner over a period of one year. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk.

Appendix Figure A3

Education, Numeracy and Risk Beliefs



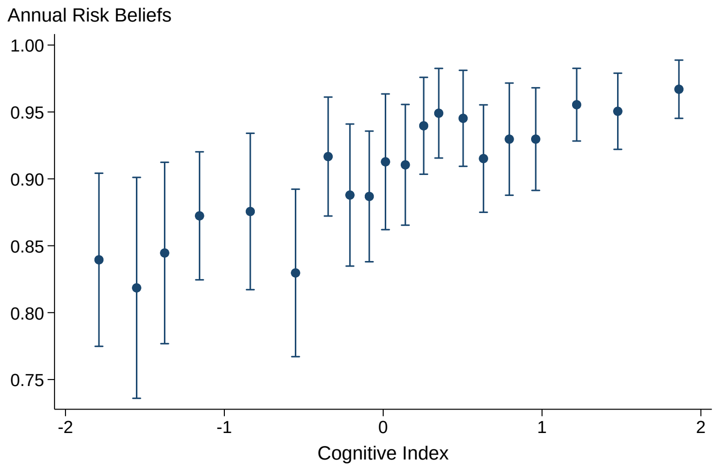
Panel A: Numeracy by Education

Panel B: Risk Beliefs by Education

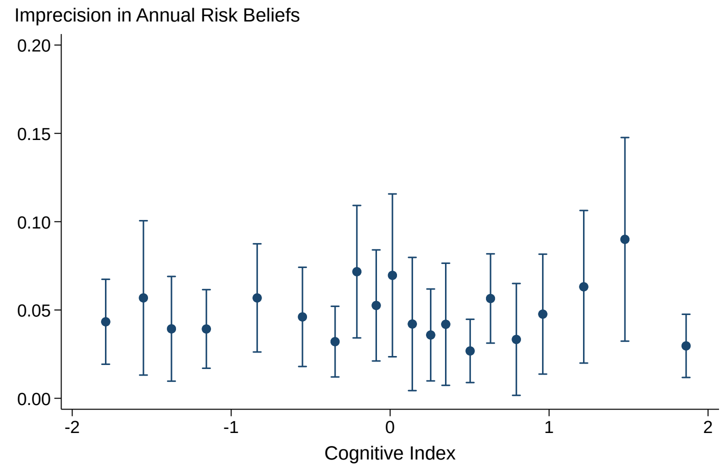
Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey. Both variables are measured at baseline. Annual risk beliefs are the perceived chance of contracting HIV from regular unprotected sex with an HIV-positive sexual partner over a period of one year. Per-act risk beliefs are the perceived chance of contracting HIV from a single unprotected sex act with an HIV-positive sexual partner. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk.

Appendix Figure A4

Risk Beliefs and Imprecision by Cognitive Index



Panel A: Annual Risk Beliefs by Cognitive Index

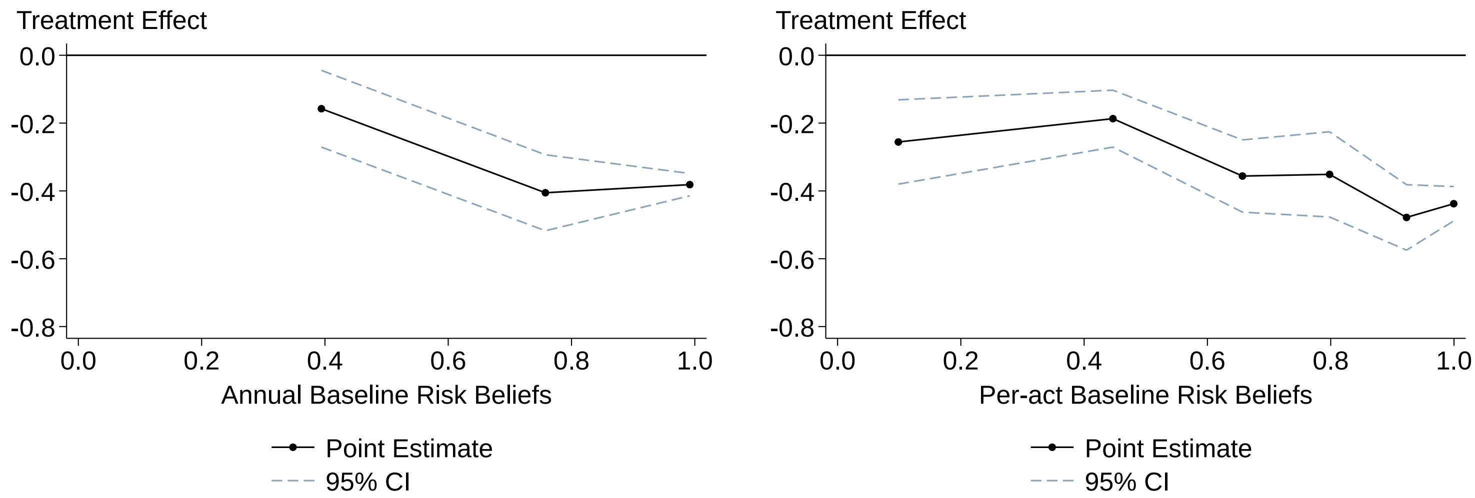


Panel B: Imprecision in Risk Beliefs by Cognitive Index

Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey. Both variables are measured at baseline. Annual risk beliefs are the perceived chance of contracting HIV from regular unprotected sex with an HIV-positive sexual partner over a period of one year. Per-act risk beliefs are the perceived chance of contracting HIV from a single unprotected sex act with an HIV-positive sexual partner. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk. The Cognitive Index is calculated using Principal Component Analysis using Years of Education, Literacy, and Raven's Score.

Appendix Figure A5

Heterogeneity in Treatment Effects on Endline Risk Beliefs
by Brackets of Baseline Risk Beliefs



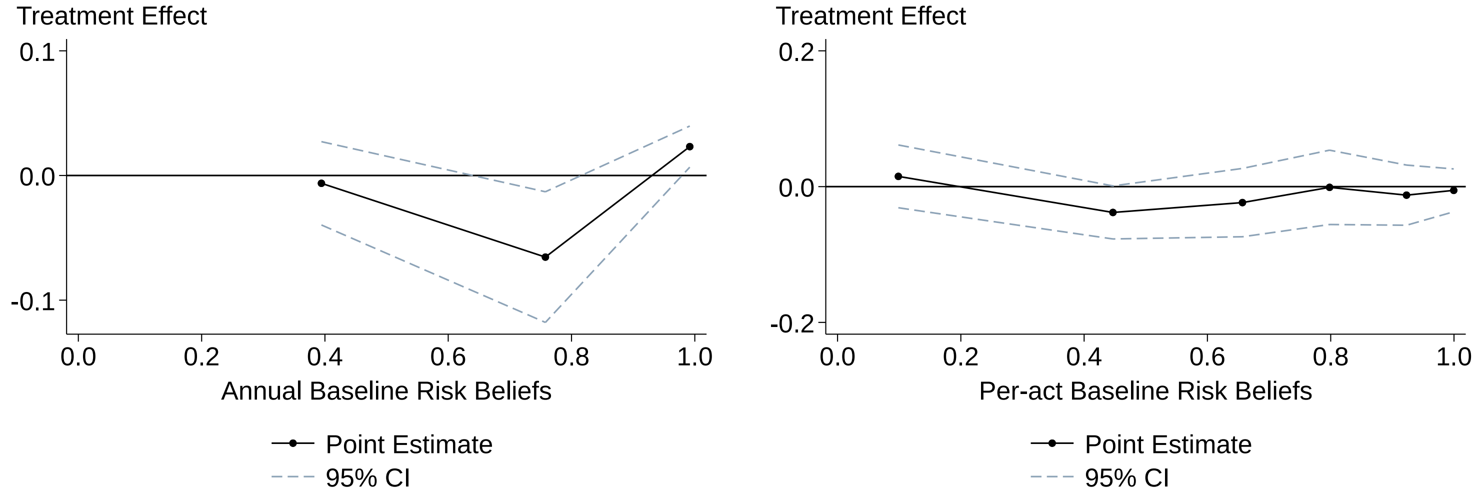
Panel A: Annual

Panel B: Per-Act

Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey. Both variables are measured at baseline. Annual risk beliefs are the perceived chance of contracting HIV from regular unprotected sex with an HIV-positive sexual partner over a period of one year. Per-act risk beliefs are the perceived chance of contracting HIV from a single unprotected sex act with an HIV-positive sexual partner. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk.

Appendix Figure A6

Heterogeneity in Treatment Effects on Endline Imprecision
by Brackets of Baseline Risk Beliefs

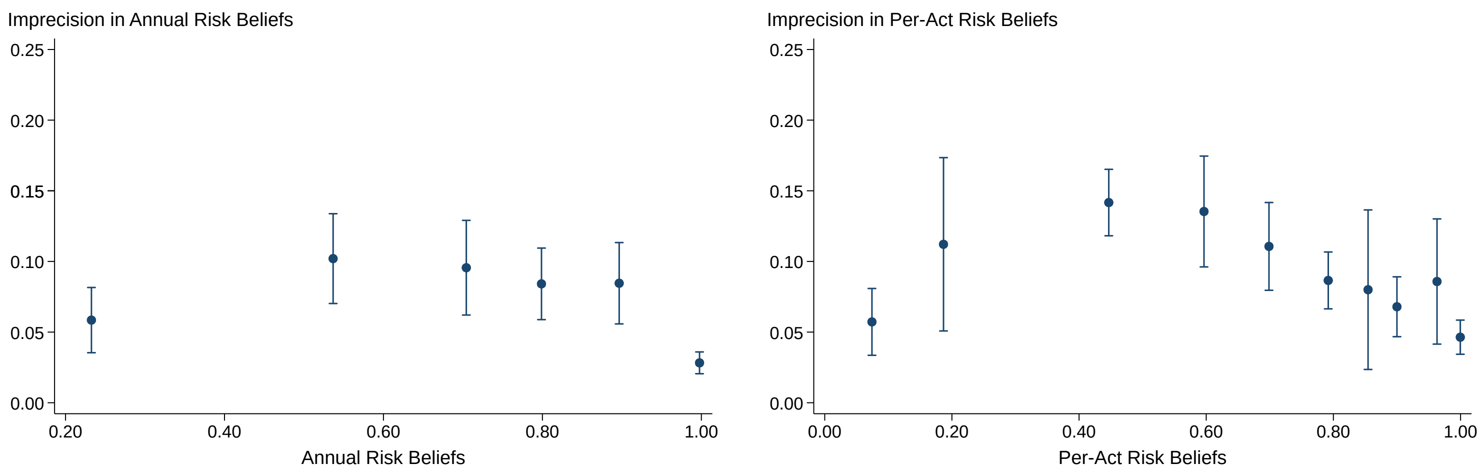


Panel A: Annual **Panel B:** Per-Act

Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey. Both variables are measured at baseline. Annual risk beliefs are the perceived chance of contracting HIV from regular unprotected sex with an HIV-positive sexual partner over a period of one year. Per-act risk beliefs are the perceived chance of contracting HIV from a single unprotected sex act with an HIV-positive sexual partner. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk.

Appendix Figure A7

Imprecision by Level of Risk Beliefs (Half-Width Instead of Range)

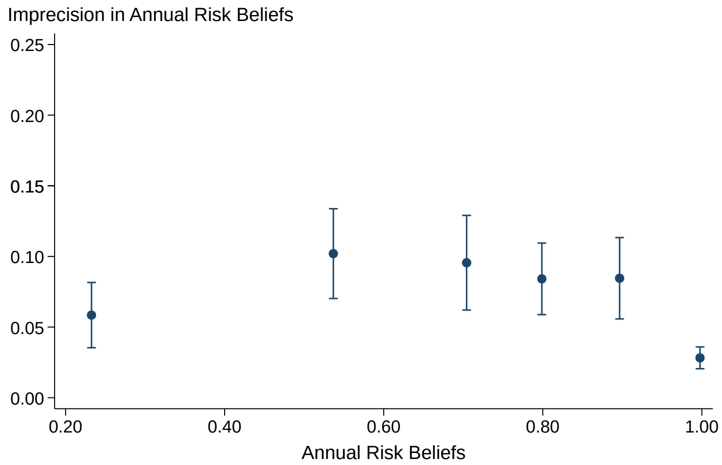


Panel A: Annual **Panel B:** Per-Act

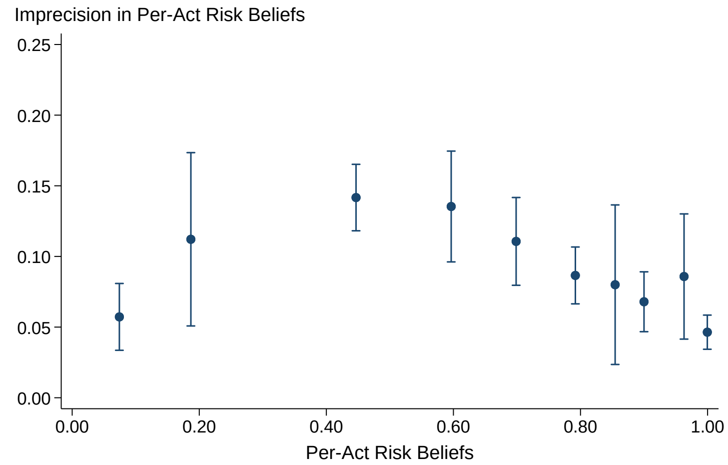
Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey. Both variables are measured at baseline. Annual risk beliefs are the perceived chance of contracting HIV from regular unprotected sex with an HIV-positive sexual partner over a period of one year. Per-act risk beliefs are the perceived chance of contracting HIV from a single unprotected sex act with an HIV-positive sexual partner. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk.

Appendix Figure A8

Imprecision by Level of Risk Beliefs (Truncated Half-Width Instead of Range)



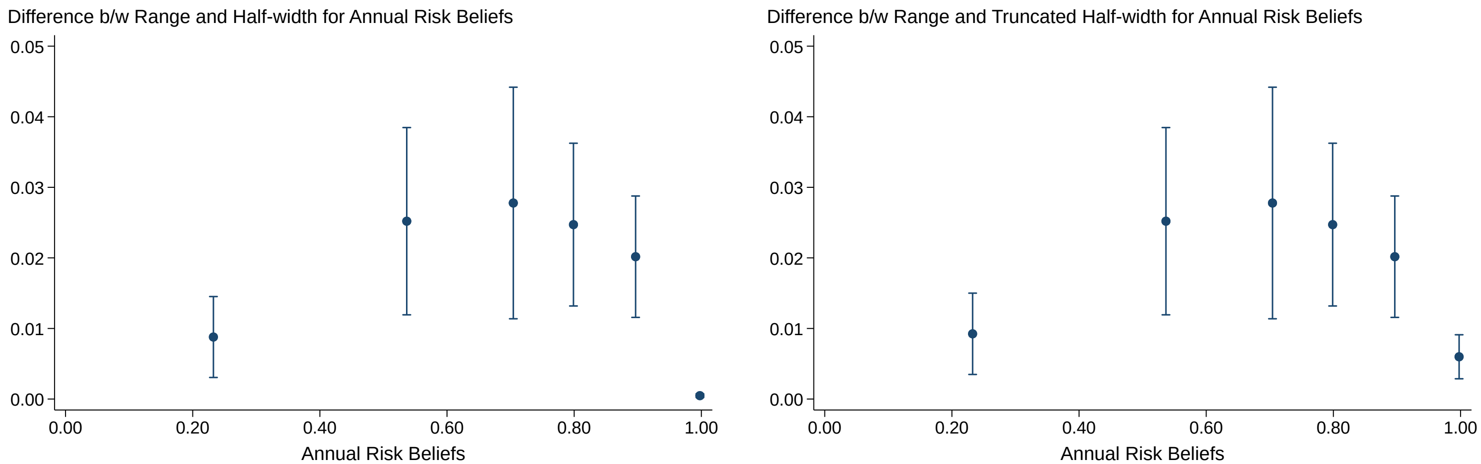
Panel A: Annual



Panel B: Per-Act

Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey. Both variables are measured at baseline. Annual risk beliefs are the perceived chance of contracting HIV from regular unprotected sex with an HIV-positive sexual partner over a period of one year. Per-act risk beliefs are the perceived chance of contracting HIV from a single unprotected sex act with an HIV-positive sexual partner. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk.

Appendix Figure A9
Difference between Imprecision Measures by Levels of Risk Beliefs



Panel A: Difference between Range and Half-width by Risk Beliefs

Panel B: Difference between Range and Half-width (Truncated) by Risk Beliefs

Notes: Sample is 1,292 people from 70 villages who completed both a baseline and an endline survey. Both variables are measured at baseline. Annual risk beliefs are the perceived chance of contracting HIV from regular unprotected sex with an HIV-positive sexual partner over a period of one year. Per-act risk beliefs are the perceived chance of contracting HIV from a single unprotected sex act with an HIV-positive sexual partner. Imprecision is measured as the width of the range that the respondent places around their point estimate of the risk.