You Know What I Know: Interviewer Knowledge Effects in Subjective Expectation Elicitation

Jason T. Kerwin and Natalia Ordaz Reynoso*

June 11, 2020

Abstract

Directly eliciting subjective beliefs is increasingly popular in social science research, but doing so via face-to-face surveys has an important downside: the interviewer's knowledge of the topic may spill over into the respondent's recorded beliefs. Using a randomized experiment that used interviewers to implement an information treatment, we show that reported beliefs are significantly shifted by interviewer knowledge. Trained interviewers primed respondents to use the exact numbers used in the training, nudging them away from higher answers; recorded responses decrease by about 0.3 standard deviations of the initial belief distribution. Furthermore, respondents with stronger priors are less affected by interviewer knowledge. We suggest corrections from the perspectives of interviewer recruitment, survey design, and experiment setup.

Keywords: Interviewer effects, Survey methodology, Subjective expectations, Risk beliefs, HIV, Spillovers

* Kerwin: Department of Applied Economics, University of Minnesota (jkerwin@umn.edu); Ordaz Reynoso: Department of Applied Economics, University of Minnesota (ordaz008@umn.edu).

Demographic research has increasingly made use of individuals' subjective expectations about probabilities and the distributions of variables. Such subjective expectations are important drivers of demographic phenomena such as fertility (Delavande 2008, Mac Dougall et al. 2013, Shapira 2013) and migration (McKenzie et al. 2013, Shrestha 2020), and can help us understand their trends and underlying determinants. Furthermore, subjective expectations are related to objective probabilities and can be used to help forecast future trends, for example in the case of mortality rates (Perozek 2008).

However, the face-to-face surveys commonly used to measure subjective beliefs in developing countries have a potential weakness: respondents' recorded beliefs may be affected by what interviewers know about the phenomenon in question. These surveys sometimes measure subjective beliefs by asking about percent chances directly (Hurd and McGarry 1995, Lillard and Willis 2001, McKenzie, Gibson, and Stillman 2006), but often use visual aids (Attanasio, Meghir, and Vera-Hernández 2005, Delavande and Kohler 2009, Delavande, Giné and McKenzie 2011a) or ask how many of a fixed number of people would have something happen to them (Aguila et al 2014, De Mel et al 2008). All three approaches rely heavily on the interviewer to explain the question and encourage the respondent to give a valid answer. These interviewer-subject interactions raise the specter of interviewer effects, and in particular the possibility that interviewer knowledge could inadvertently spill over onto subjects' recorded beliefs.

The effect of interviewer characteristics on survey responses has been documented across a wide range of contexts. Examples of these characteristics are race and ethnicity (Cotter et al. 1982, Reese et al 1986, Anderson et al. 1988, Finkel et al. 1991, Davis 1997, Dionne 2014, Adida et al. 2016), religion (Blaydes and Gillum 2013), gender (Becker, Feyisetan and Makinwa-Adebusoye 1995, McCombie and Anarfi 2002), and social or cultural proximity (Weinreb 2006). Respondents may also infer the purpose of the study from interviewers and change their answers as a result, a pattern known as experimenter demand effects (Orne 1962, Zizzo 2010, de Quidt et al. 2018). This body of research shows the importance of social interactions in the interview setting for recorded survey responses, and how interviewer characteristics may impact this interaction. An extensive literature has also explored the methodology of subjective belief elicitation (Delavande 2014). However, to our knowledge, no previous paper has studied the role of interviewer knowledge in driving survey responses.

Leveraging a randomized experiment that used interviewers to implement an information treatment, we show that interviewer knowledge has an effect on respondents' recorded beliefs. The experiment was designed to investigate how information about the true transmission rate of HIV affects risk-taking (Kerwin 2018). Interviewers were taught the true HIV transmission rate midway through the baseline survey, in order to conduct the information treatments for the study. The study respondents were randomly divided into control surveys, which happened before the interviewers learned the information, and treatment surveys, which happened afterwards. We use data from the baseline surveys, when treatment-group respondents had not yet been taught the risk information themselves, but were interviewed by people who had been taught it.

Interviewer knowledge matters for recorded risk perceptions. Comparing the baseline surveys across study arms, we find that interviewers who were exposed to the information treatment elicit lower HIV transmission rate perceptions from respondents. Reported beliefs are significantly shifted by the interviewers' knowledge, decreasing by about nine percentage points, or roughly 0.3 SD of the control-group belief distribution. This finding can help explain the puzzling finding that people's preferences and beliefs appear to be very unstable in panel surveys (Chuang and Schecter 2015, Mueller et al. 2019). If recorded responses are heavily shaped by interviewers' knowledge and beliefs, then people's answers may appear to be much more unstable than they really are.

In addition to shedding light on the role of interviewer knowledge in driving survey responses, our study also builds on the previous literature on interviewer effects by isolating the causal effect of a specific interviewer characteristic—knowledge. Past studies of interviewer effects have been able to exploit the exogenous assignment of interviewers to respondents, but have been limited by the fact that the interviewer characteristics in question are both fixed and correlated with other attributes. For example, race is correlated with income and socioeconomic status, and a wide range of interviewer characteristics can all affect responses simultaneously (Di Maio and Fiala 2019). Because interviewers in our study were exogenously shocked with new information about HIV transmission rates, we can isolate the causal effect of knowledge alone. This is the first study we are aware of that has been able to identify the causal effect of a single interviewer characteristic. This is possible because knowledge, unlike the other characteristics that are typically studied, is malleable: it can be changed quickly, whereas even many non-fixed traits like education levels can be changed only slowly, and others such as age cannot be changed at all.

We can identify several channels through which interviewers' knowledge affects recorded risk perceptions. First, interviewers who underwent the training primed respondents to give answers that match the exact training content. The training explained that the annual transmission rate of HIV between an HIV-positive spouse and an HIV-negative spouse who have regular unprotected sex is 10%. Consistent with a priming story, treatment-group respondents are 4.3 percentage points more likely to (incorrectly) report that the *per-act* probability of HIV transmission is exactly 10%.

A related mechanism by which interviewer knowledge affects recorded risk perceptions is through nudging respondents to give lower answers. Evidence for this comes from an aspect of the survey design: if a respondent answered exactly 50% for any risk perception question, interviewers were taught to follow up and see if they were simply unsure; if so, they were asked for their best guess, following Hudomiet, Kézdi, and Willis (2011). Interviewers who underwent the training were less likely to elicit higher numbers when they ask respondents to provide a best guess in this situation. This suggests that interviewers who have been exposed to the information treatment are nudging participants away from higher answers. The same pattern could also affect the initial responses to the questions.

The strength of respondents' priors may affect how much interviewer knowledge matters. The effect of interviewer training is smaller for more educated respondents, and falls to zero for respondents who reached at least Form 2 (10^{th} grade) in school. This may be due to the fact that students in Malawi learn about HIV transmission during Form 2, and are exposed to a narrative that claims HIV is highly contagious. While the information taught during Form 2 diffuses through the population as a whole, more-educated respondents are exposed to it directly, and thus likely feel more certain about their beliefs. This makes them less susceptible to the interviewer's nudges to report lower risk beliefs.

We suggest several ways to correct for interviewer knowledge effects. Interviewer recruitment for face-to-face surveys should try to match the population of respondents, and interviewer training should emphasize the possibility of unintentional spillovers and the need to treat all respondents consistently. When designing information experiments, researchers should consider running baseline surveys simultaneously across groups or separating the information treatment from surveys, although these approaches have their own drawbacks. Even when no information is provided, knowledge spillovers remain likely—interviewers can vary widely in their knowledge and beliefs, and so teaching them about topics relevant to the survey may improve data quality. One promising avenue is to eliminate the interaction between interviewer and respondent by performing surveys via audio computer assisted self interviewing (ACASI) which can work even in contexts with low literacy and numeracy. However, ACASI may lead to lower data quality than face-to-face interviews; further work in this area would be invaluable.

The remainder of this paper proceeds as follows. Section 1 reviews the role of subjective beliefs in demography, provides background on Malawi's HIV epidemic, and describes the previous literature on subjective expectations about HIV in Malawi. Section 2 describes the experiment and the data, and Section 3 presents the empirical strategy we use to evaluate interviewer knowledge effects. Section 4 shows how interviewer knowledge affects measurements of respondents' subjective expectations. In Section 5 we explore mechanisms for the effects, and in Section 6 we discuss potential corrections for interviewer knowledge effects. Section 7 concludes.

1 Background

1.1 Subjective Expectations and Demographic Decisions

Subjective expectations play a key role in driving demographic patterns and people's responses to them. They drive contraception choices and fertility (Delavande 2008, Mac Dougall et al. 2013, Shapira 2017) as well as migration (McKenzie et al. 2013, Shrestha 2020). Perceived mortality risks affect whether people engage in life-threatening activities—see e.g. Delavande and Kohler (2016) for HIV, León and Miguel (2017) for transportation choices, and Bennear et al. (2013) and Keskin et al. (2017) for water safety—and thus actual mortality rates. The effects of subjective expectations often spill over between demographic choices and other domains. For example, women's education and career decisions depend on their beliefs about the costs of raising children—which can differ sharply from reality (Kuziemko et al. 2018). Similarly, women tend to systematically underestimate their fecundity at young ages and overestimate it at older ages (Mahony 2011).[1]

[1] In addition to uncertainty about their material circumstances, people may also be uncertain about their own futures or the correct course of action; periods of uncertainty can cause intersections between aspects of people's lives that usually seem separate, such as work and romantic relationships (Johnson-Hanks 2017). This uncertainty about the correct course of action is itself shaped by uncertainty about material facts, such as the risks of mortality and miscarriage (Trinitapoli and Yeatman 2018).

Data on people's subjective expectations are also a potentially-useful tool for demography research. Subjective mortality probabilities may be useful predictors of actual mortality rates (Perozek 2008) and are correlated with known predictors of mortality (Delavande et al. 2017), although there is also evidence that people's subjective mortality beliefs have systematic biases (Elder 2013). Despite their limitations, subjective mortality beliefs may still be valuable: people form them using risk factors such as parental health and longevity that objective mortality rates cannot account for, and they affect risk-taking behaviors (Dormont et al. 2018). Similarly, self-rated health is a useful predictor of mortality (Burström and Fredlund 2001). Subjective beliefs about health can help forecast future mortality rates: they are available earlier than objective predictors of mortality, and they predict mortality even conditional on objective measures of health status (Idler and Benyamini 1997).

There is a crucial distinction between individuals' subjective expectations about risks and other variables and the true values of these figures. Much social science research assumes that people know the true values of numbers, but recent research has focused on measuring people's actual perceptions, which can be quite different from the truth (Manski 2004). Consider the case of subjective expectations about mortality rates. These can differ from true population-average mortality rates in three key ways (Delavande and Rohwedder 2011). First, they measure a variable that has not yet been observed because the population answering the survey questions is still alive. Second, they may be measured with error relative to the person's true beliefs. Third, they reflect individuals' beliefs about what will happen, rather than the truth.

1.2 The HIV Epidemic in Malawi

Malawi has been dealing with a severe HIV epidemic for several decades and the disease has had major effects on its population. The prevalence of the virus has been stable at around 10% of the population for roughly the past decade (NSO Malawi and ICF 2017).[2] The expansion of access to antiretroviral treatment (ART) for HIV has drastically improved the situation for HIV-positive people in recent years. Starting in 2016, Malawi implemented a universal test-and-treat policy, so that all HIV-positive people had access to ART (Alhaj et al. 2019). Testing rates are still

[2] A small apparent drop (to a prevalence of 9%) in the 2015 DHS was the result of a change in methodology; measured on a consistent basis, the prevalence was essentially unchanged from the 2011 survey.

low for men, but most women get access to treatment because there is strong pressure to accept the nominally-voluntary HIV tests during antenatal care visits (Angotti et al. 2011).[3] Even with the expansion of access to treatment, however, HIV is still a major issue in people's lives.

The large scale of Malawi's HIV epidemic has led to extensive research by social scientists on how it impacts people's lives. Most prominently, this includes the Malawi Longitudinal Study of Families and Health (MLSFH), which has been collecting demographic, socioeconomic, and health information on the same households since 1998. The MLSFH also embeds a novel ethnographic study, the Malawi Journals Project (MJP), in which Malawians record everyday conversations about HIV/AIDS.

1.3 Subjective Expectations about HIV in Malawi

An important focus of research on HIV in Malawi has been on measuring people's subjective beliefs about the disease and understanding how those beliefs affect their behavior. The MLSFH measures both people's perceptions about HIV and their sexual activity, and has an embedded experiment in which respondents were incentivized to learn their HIV status (Thornton 2008, Fedor et al. 2015). It was also used as a platform to design and study an innovative technique for capturing subjective probabilities using visual aids (Delavande and Kohler 2009). This work was an important contribution to a literature that shows that eliciting subjective probability beliefs is feasible in low and middle-income settings (Delavande, Giné and McKenzie 2011b, Delavande 2014, Attanasio 2009).

A core finding of the work on subjective beliefs about HIV in Malawi is that people substantially over-estimate their likelihood of being HIV-positive (Bignami-Van Assche et al. 2007, Anglewicz and Kohler 2009). Relatedly, they also over-estimate the transmission rate of the virus by several orders of magnitude (Delavande and Kohler 2016, Kerwin 2018). Extensive research has tried to understand how people form these beliefs. One channel is through HIV testing: Malawians who learn they are HIV-positive lower their beliefs about the transmission rate of the virus (Delavande and Kohler 2012); this may be because they realize they have not yet transmitted the virus to their sex partners. Qualitative evidence from the MJP supports this

[3] This is part of a systematic pattern of HIV prevention efforts targeting women and excluding men (Watkins 2011).

quantitative finding. Kaler and Watkins (2010) find that people are ambivalent about testing: they think that it will always lead to a positive result, followed by death. They also find that as a result of thinking that HIV tests mostly turn out positive, people overestimate the transmission rate of HIV.

Information from HIV tests spreads beyond the person being tested. Spouses typically tell each other about their HIV test results, although HIV-positive women are less likely to share their status (Anglewicz and Chintsanya 2011). More broadly, subjective expectations about HIV risks spread through social networks (Helleringer and Kohler 2005, Kohler et al. 2007). People also draw inferences about HIV risks from their own experiences. For example, when young women marry, they are more likely to think they are at risk of contracting HIV in the future—possibly because they know or suspect their husbands are unfaithful (Grant and Soler-Hampejsek 2014).

Another line of research has shown that subjective expectations about HIV matter: people respond to their perceived risks of being HIV-positive. It is perceived, rather than actual, HIV status that drives condom use for women, even when one's actual status is known (Anglewicz and Clark 2013). Ethnographic evidence from the MJP shows a similar pattern for men. They assume they are HIV-positive even with no medical evaluation or signs of AIDS, which drives further risky behavior (Kaler 2003, Kaler et al. 2015). Many people are uncertain about their HIV status, and this uncertainty affects their fertility intentions (Trinitapoli and Yeatman 2011).

In addition to changing their behavior in response to their perceived HIV status, people also respond to their perceived chance of contracting the disease. Grant and Soler-Hampejsek (2014) show that women may use divorce to protect themselves if they believe they are at high risk of contracting HIV, mirroring the finding by Anglewicz and Reniers (2014) that HIV-positive people have higher rates of widowhood and divorce. Women who anticipate that they will contract HIV in the future invest more in their children's education (Grant 2008). The causal effect of risk perceptions on behavior also holds for probabilistic beliefs of the kind studied in this paper (Delavande and Kohler 2016, Kerwin 2018).

2 Data and Empirical Design

We use data from an experiment designed to study the effects of risk perceptions on risk-taking behavior (Kerwin 2018) that was conducted in the Zomba District of Malawi from August to December 2012. The experiment randomly assigned half of respondents (stratified by distance to the trading center and gender) to receive information about HIV transmission risks at the end of the baseline survey. Treatment-group participants were read an information script that explained the actual HIV transmission rate for couples with one infected partner that have regular unprotected sex (on average about 100 times per year). The true transmission rate is 10% per year (Wawer et al. 2005), far below what Malawians typically believe. In our sample, the average risk belief is about 90% per year, and nearly half of our sample thinks the transmission rate from just a single exposure is 100%.

The risk information was provided by the survey interviewers, using a script and a set of visual aids that were built into the treatment-group surveys. The interviewers themselves were taught the risk information and how to conduct that survey module via a two-day training session that took place halfway through the baseline data collection. All the control-group surveys were scheduled to occur before this training session to minimize the risk of contaminating the control villages, following Godlonton, Munthali, and Thornton (2015).[4]

The interviewers seem to have been unaware of the actual HIV transmission rate prior to the training session, and thus it likely strongly shifted their beliefs about HIV risks. While we lack direct data on their beliefs prior to the information session, two sources of evidence support this claim. First, the interviewers all lived in or close to the study area, so the pre-training data for the control group is a reasonable proxy for their beliefs. Less than two percent of our control group thought the annual risk of HIV transmission was below 20% at baseline. A second piece of evidence comes from the training session itself. The interviewers expressed surprise at the information they were taught, and many were initially reluctant to believe it. To help convince them, project staff had to show them the original research study (Wawer et al. 2005) as well as the section of the Malawi National AIDS Commission website that listed the HIV transmission rate.

The fact that the interviewers were taught new information just before they started to survey the randomly-assigned treatment group allows us to study how that information affected survey responses. We use the interviewer training session as a treatment, and study how that

[4] There were a handful of control surveys that did take place after the training session. These were cleanup surveys that happened because selected respondents were not available at the time of the scheduled baseline interview, and had to be interviewed afterwards.

changes the recorded baseline beliefs of respondents. Comparing the baseline beliefs between the treatment and control groups allows us to estimate the effect of interviewer knowledge on recorded risk beliefs. Online Appendix Figure 1 shows the timeline of the experiment.

Our principal outcome measure is respondents' recorded subjective risk beliefs on the baseline surveys. This variable was collected by asking questions about proportions out of a fixed number of people. For example, "If 100 men, who do not have HIV, each have sex with a woman who is HIV-positive tonight and do not use a condom, how many of them do you think will have HIV after the night?" The questions cover per-act and per-year transmission rates for both protected and unprotected sex. Respondents picked integers between 0 and 100 in response to each question.[5] The exact wording of all four questions is in Online Appendix Figure 2. This style of expectation question has also been tested and validated by previous research in Malawi (Chinkhumba, Godlonton, and Thornton 2014, Godlonton, Munthali, and Thornton 2015, Kerwin et al. 2011).[6] Interviewers had no incentive to record specific answers to this question, but instead were incentivized to record answers accurately: random back-checks were used to check that surveys actually happened and responses were written down correctly.

Our sample of respondents includes 1,292 individuals from 70 villages who have both valid baseline and endline survey data. Baseline demographic statistics for the treatment and control groups can be found in Online Appendix Table 1; the two study arms were balanced on observable exogenous variables. The experiment we use was not designed to study the interviewers, and so we have very limited data on their characteristics based on administrative records. There are 14 total in our sample; half were female and half were male. The even gender split was chosen intentionally to facilitate gender-matched interviews: all male respondents were interviewed by men and all women were interviewed by women. All of them had completed secondary school (a screening requirement imposed during hiring) and most had graduated recently (so were in their 20s). They were recruited from the local area, but were not assigned to survey anyone they knew

[5] For the unprotected per-year question, the correct answer is 10; for the unprotected per-act version, the closest possible answer to the truth is 0. In the absence of condom failures, the correct answer for both the per-year and per-act condom-protected questions is 0.

[6] These questions measure the respondents' perceived risk of contracting HIV from various sexual behaviors—not their perceived probability of currently being HIV positive. Our measured probabilities are comparable to other measurements from Malawi. For example, in Delavande and Kohler (2012) the mean unprotected per-act risk is 87%; in our data it is 83%.

personally.

3 Empirical Strategy

To study the effect of interviewer knowledge on respondents' recorded risk beliefs, we compare the baseline recorded beliefs of the treatment and control groups. Our main regression specification is Equation 1, where Y_i is either a measure of risk belief at baseline, or an indicator variable of specific values of the risk belief at baseline. The dummy variable T_i takes a value of one for respondents in the treatment group and zero otherwise. Our treatment is thus defined as having been interviewed at baseline by a more-knowledgeable interviewer. We control for sampling strata fixed effects Z_i and interviewer fixed effects I_i; the latter allow us to rule out the possibility that other interviewer characteristics, besides knowledge, are driving our results. We also control for W_i, a sexual activity index based on the first five variables in the balance table (see Section 4.2 for further discussion). All standard errors are adjusted for clustering by village.

$$Y_i = \alpha + \beta T_i + Z_i'\eta + I_i'\gamma + \delta W_i + \varepsilon_i \tag{1}$$

To understand the mechanisms behind the effects, we interact the treatment indicator with respondent characteristics (Equation 2). We de-mean all covariates before interacting them with the treatment indicator, so the main effect of the treatment can still be interpreted as the sample-average treatment effect (Imbens and Rubin 2015, p. 247).

$$Y_i = \alpha + \beta T_i + \gamma\, T_i * X_i + \delta X_i + Z_i'\eta + I_i'\gamma + \delta W_i + \varepsilon_i \tag{2}$$

4 Results

4.1 Main Estimates

Interviewers exposed to the information treatment elicit lower risk perceptions. Figure 1 shows the daily average recorded risk beliefs for the treatment and control groups at over time. The first group of observations represents the baseline control group beliefs, when neither the interviewers nor the respondents knew the content of the information treatment. After those surveys were conducted, the interviewers learned the content of the information treatment (vertical dashed line), and then did the baseline treatment surveys. We can see that the treatment-group beliefs are lower than the control-group beliefs.

There are five days with control-group baseline data after the information treatment. These are for cleanup baseline surveys for the control group that were conducted after the bulk of the baseline control-group surveys were finished, and took place after the interviewer training session. This happened when respondents were not available at the initially-scheduled baseline interview. The distribution of beliefs for these observations is closer to that of the treatment group than to the rest of the control group. This lends support to the idea that it is interviewer knowledge specifically, and not some other factor that is imbalanced across study arms, that is causing the mean difference between baseline treatment and control recorded beliefs.

Further support for the idea that the change in beliefs is due to interviewer knowledge is evident in the endline risk beliefs. First, the endline risk beliefs allow us to reject the possibility that the treatment group simply accidentally received the risk information prior to answering their baseline survey questions. The direct information treatment effect on risk beliefs (the gap between the endline risk beliefs for the treatment and the control groups) is much larger than the treatment-control difference at baseline.

Second, the control-group endline beliefs are very similar to the treatment-group baseline beliefs. This is completely consistent with a model in which recorded beliefs are moved by interviewer knowledge: neither the treatment group at baseline nor the control group at endline had been directly told the information about HIV transmission risks, but both were interviewed by interviewers who did know the information. As a result, both sets of beliefs are shifted downward relative to the control-group baseline beliefs, and they also have similar average values to one another.

Table 1 presents our main results numerically. Each column represents a measure of a different HIV transmission risk: per act or per year, using condoms or unprotected. For all four measured risk beliefs, the coefficient of the treatment (interviewer training) is negative and significant. In the case of the per-act unprotected transmission risk, the coefficient is 9.3 percentage points, or about 0.3 standard deviations. The magnitude of the effect is relatively large, especially considering that it is an unintentional spillover: respondents were not directly exposed to the information treatment. As can be seen in Figure 1, the effect at end-line, when participants themselves were exposed to the information treatment, was larger: 38.4 percentage points for the perceived per-act, unprotected transmission risk.

Participants in the control group had average beliefs that were substantially larger than the true rate of HIV transmission in each one of those cases. For example, the true value of the per year, unprotected transmission rate is about 10% (Wawer et al. 2005), but the average respondent in the control group thought the risk was 83%, and well over half of respondents thought the risk was 100%. Baseline beliefs for the control group have the correct ordering in terms of which risk is higher, but the average levels are higher than all true infection risks.

Interviewer training decreased recorded risk beliefs for all four measures, even though the training only discussed the unprotected per-year risk, shown in Column 2. Column 1 shows an effect of 9.3 percentage points (0.35SD), columns 2 and 4 show effects of 4.8 and 7.9 percentage points respectively (0.28SD each). Column 3 shows the smallest effect: 2.7 percentage points (0.12 SD), corresponding to the per-act, condom-protected transmission risk.[7] This variable has the lowest control group mean overall, so a smaller effect is not surprising, and we can still reject a zero treatment effect. Moreover, condom protected risks are simply scaled-down versions of the unprotected risks, so changes in those variables should be smaller.

The fact that interviewer knowledge changes responses for risk beliefs that were not explicitly targeted speaks to interviewers internalizing the information and actually changing their beliefs about transmission risk, as opposed to memorizing the one figure that was presented to them. Interviewers know that the four measures of transmission risk are related, and when they adjust their beliefs for one, this impacts their beliefs of all others. This makes the threat of interviewer knowledge effects more general, as knowledge spillovers may occur not only with directly-provided information but with its implications as well.

4.2 Alternative Explanations

Interviewer Experience

Another potential explanation for our findings is interviewer experience. The trajectory of the pre-treatment trend in the first portion of Figure 1, if extended, would intersect the level of

[7] The results in Table 1 are qualitatively identical if we include interviewer fixed effects as controls (Online Appendix Table 2).

recorded risk beliefs in the second portion. This could have happened if interviewers improved over time at asking the relatively complicated questions on the subjective expectation module. Our basic results in Table 1 do not rule out the possibility that the estimated treatment effects are due to interviewer experience alone.

To examine this possibility in further detail, we present a set of regression discontinuity (RD) plots in Figure 2. These graphs are produced using the Calonico et al. (2015) `rdplot` Stata command to automatically bin the data and fit polynomial curves on either side of the discontinuity. The binned averages are shown using dark gray dots, with black lines for the polynomial fits. The light gray regions show 95% confidence intervals for the bin-specific averages. We also show the estimated treatment effects from regression discontinuity models in Table 2, using the `rdrobust` Stata command (Calonico et al. 2017). This command automatically selects data-driven bandwidths and computes robust bias-corrected p-values (Calonico et al. 2014).

Panel A of Figure 2 shows the main comparison of interest: before vs. after the training session, for the unprotected per-act HIV transmission risk belief. Two conclusions are clear from the graph. First, the steep downward trend apparent in the first portion of Figure 1 was partly an artifact of fitting a linear trend to daily average risk beliefs, as opposed to the underlying survey data. Fitting a flexible polynomial to the actual survey responses reveals a slight downward trend to the left of the discontinuity. There is some evidence for interviewer experience driving a downward trend in responses, but the pattern is not particularly strong.

Second, even accounting for trends in responses due to interviewer experience, there is a sharp jump in responses right at the time of the intervention. The polynomial fits differ by over 10 percentage points, and the bin-average confidence intervals barely overlap. The numerical results (Column 1 of Table 2, Panel A) show that this jump is statistically significant: the RD estimate of the treatment effect is 15 percentage points, with a p-value of 0.03.

Another way of assessing the role of interviewer experience is to compare the results to another complex survey module. The questions about sexual activity in the past week were collected using a retrospective sex "diary" that was originally developed by Kerwin et al. (2011). This module had interviewers walk respondents through each of the previous seven days to record details about each sex act on each day as well as other events on that day. The other events included when they woke up and went to sleep and whether they or their partners were menstruating, as well as alcohol consumption. These other details were included to capture risk factors, and also to help respondents remember specifics about their sexual activity, similar to an event history calendar (Belli et al. 2001) or a relationship history calendar (Luke et al. 2011). This module was complicated to carry out and required the most attention when teaching interviewers to conduct the survey. If the complexity of the HIV risk perception questions led to changing response patterns as interviewers gained experience, we might also expect a similar pattern for the sex "diary" questions.

Panel B of Figure 2 presents an RD plot for the number of sex acts in the past week as reported on the sex "diary" module. There is no clear trend for most of the left-hand side of the graph (before the HIV information training session), although there is a dip visible just before the training session. Notably, this dip is matched on the right-hand side of the graph, so that the confidence intervals for the bins just before and just after the discontinuity largely overlap. In Column 2 of Table 2, Panel A, we show numeric estimates of the size of this regression discontinuity. Consistent with the graph, there is a small positive but statistically-insignificant difference.

The HIV information training session occurred during a six-day gap in the data collection schedule. Could something about a break in surveying be creating the differences in the baseline responses? To assess this, Panel C of Figure 2 looks for discontinuities in responses between the end of the baseline survey and the beginning of the endline survey; between the two survey waves there was a ten-day break in data collection. Using the control-group data only, it plots an RD for the recorded HIV risk perceptions in the baseline surveys (left side of the graph) vs. the endline surveys (right side). The confidence intervals overlap, and the estimated difference (Column 1 of Table 2, Panel B) is nearly zero and statistically insignificant. We see similar null results for the number of sex acts in the past week from the sex "diary". This suggests that a break in surveying does not appear to have *per se* effects on the recorded survey responses.

A potential threat to the identification of these RD estimates is that there could have been systematic sorting of respondents around the breaks in data collection. If different kinds of respondents were interviewed just before the training session vs. just after, it would be incorrect to attribute the 15 percentage-point drop in recorded risk beliefs to the effect of the training session. To test for this sort of systematic sorting, Columns 3 and 4 of Table 2 present RD estimates for fixed respondent characteristics: gender and age. There are no statistically-significant differences in either characteristic for the HIV information training session nor for the end of the baseline.

Imbalance

A second potential explanation for the differences between the treatment and control-group beliefs at baseline is imbalance. Although our randomized experiment ensures that the two groups were balanced in expectation, for any given realization of the random assignment it is possible for them to have differences (Frison and Pocock 1992). Those differences could in turn lead to different beliefs. A particular concern is balance on sexual activity, which is correlated with risk beliefs (Smith and Watkins 2005). While the sexual activity variables in Panel A of Online Appendix Table 1 are balanced overall, the first five rows all show higher values for the control group than the treatment group. To test for an aggregate balance problem in these variables, we construct an alternate sexual activity index that uses those first five variables alone. The difference is not statistically significant ($p = 0.149$). However, even a statistically-insignificant difference in this variable could lead to substantively-important differences in the belief variables. To mitigate this concern, we control for this alternate sexual activity index in all our regression analyses. Our results do not depend on this choice: the main effects on beliefs from Table 1 are barely changed if we drop this control (Online Appendix Table 2) or if we drop the interviewer fixed effects as well (Online Appendix Table 3).

Another potential source of imbalance is variations in religion, ethnicity, and languages spoken across the two groups. HIV risk perceptions and sexual behavior vary widely by religious denomination in Malawi (Trinitapoli and Regnerus 2006, Trinitapoli 2009) and ethnicity-specific cultural activities such as initiation rites are ways that people learn about sexual health (Munthali and Zulu 2007). Administering surveys in an unfamiliar language can lead to item non-response and systematic measurement error (Andreenkova 2018). This could be an issue since all our surveys were administered in Chichewa, but this concern is substantially mitigated by the fact that virtually all of our subjects are fluent speakers of either Chichewa or the mutually-intelligible language Chinyanja. In the 1998 Malawi census, 96% of households in the study area (TA Mwambo) reported that their most-commonly used language was Chichewa or Chinyanja (Minnesota Population Center 2019).

Online Appendix Table 4 shows balance statistics for specific religious denominations as well as ethnic groups. Panel A shows that while the treatment is balanced in terms of the share of Christians and Muslims (Online Appendix Table 1), there are important differences across study arms for some specific denominations. In contrast, the treatment is fairly balanced by ethnic group

(Panel B). However, this pattern may mask potential differences in language abilities within ethnic groups. The survey did not ask respondents whether they speak Chichewa at home or whether it was their first language, so we use Chichewa-language literacy as a proxy. Online Appendix Table 5 shows balance statistics for people being literate in Chichewa by ethnic group. There are no large differences, but the two percentage-point difference for the "other" group is statistically-significant.

To account for potential differences in responses driven by these variations in religion and ethnicity, we add indicators for membership in each group to our regression. We also add indicators for a person being from a given ethnic group and also literate in Chichewa. The results are in Online Appendix Table 6. The effects on measured risk beliefs are essentially unchanged: they remain statistically significant and are slightly larger on average.

Spillovers

The similarity in responses between the treatment baseline and control endline surveys implies that interviewer knowledge drives our results, rather than some other change that occurred at the time of the training session. This similarity could also have arisen through spillovers: if treatment-group respondents told control-group respondents about the information they learned, then we would expect a fall in control-group beliefs. To test for this possibility, we use social network data to count the number of total friends each respondent has, and the number they have that live in treatment-group villages. We then estimate

$$Y_i = \alpha + \beta T_i + \eta TotalFriends_i + \gamma TreatedFriends_i + \varepsilon_i \tag{3}$$

where Y_i is the respondent's endline risk belief; we also run versions of the regression that break out the spillovers by study arm. This regression identifies spillover effects on endline beliefs, since a respondent's number of treated friends is randomly assigned conditional on their total number of friends (Kremer and Miguel 2007). The results are in Online Appendix Table 7. We see no evidence of spillovers onto the control group. Another possible kind of spillover is that control-group respondents might have sought out information about HIV because they were asked about it. We cannot rule out this possibility, but it is unlikely to have generated the observed empirical pattern. This information seeking would have had to lead to endline beliefs that are nearly identical to the (measured) baseline beliefs for the treatment group, who did not have any time to seek out information about HIV risks prior to answering the risk belief questions at baseline.

5 Mechanisms

Our results show that being surveyed by a more-knowledgeable interviewer causes a decrease in recorded risk beliefs, and that this effect occurs not just for the beliefs that the interviewer was directly taught about, but also for other related risks. How do these spillovers between interviewer beliefs and the (recorded) beliefs of survey respondents happen? We explore several possible mechanisms for the effect.

5.1 Priming

Since one-on-one surveys involve a face-to-face conversation between respondents and interviewers, it is possible that interviewer knowledge could affect recorded responses via priming. We find evidence that trained interviewers primed respondents to give answers that matched up with the exact numbers used in the training. Table 3 shows regressions of indicator variables that take a value of one when respondents answer exactly 10% for each one of the risk belief questions. Ten percent is the exact figure that the interviewer training provided as the true value of the per year unprotected HIV transmission risk. Interviewer training makes respondents more likely to answer exactly 10% for the per-act unprotected risk (column 1), even though that is not the true risk.[8] We therefore interpret this coefficient as the result of interviewers priming or nudging respondents towards lower responses to all risk belief questions, not just the one corresponding to the information treatment. However, we do not see an increase in reporting an answer of exactly 10% in Column 2 (the annual risk), where it is the correct answer. A potential explanation is that respondents have extremely high priors for this figure: the average risk belief is 93%. In columns 3 and 4 (condom-protected risks) we see slight reductions in the chance that people report exactly 10%. Since those questions immediately followed the unprotected risk questions, this could be explained by respondents updating their risk beliefs in a consistent way: if condoms lower the risk by a factor X, and the unprotected risk is 0.1, then the condom-protected risk is 0.1X.

These results are consistent with the literature on priming and anchoring, which shows that

[8] The information treatment only mentioned the annual unprotected transmission risk, and the figure provided for the true risk was 10%. The true per-act transmission risk is approximately 0.1%.

mentioning numbers will induce people to give answers to subsequent questions that are more similar to those numbers (Newell and Shanks 2014). This can happen by directly suggesting a potential answer, exposing respondents to peers' responses (Tversky and Kahneman 1974), or even by mentioning totally unrelated numbers (Chapman and Johnson 2002, Mussweiler et al. 2000).[9] While all three priming pathways are possible in our context, the first is the most likely. Interviewers were trained to encourage respondents to answer even if they were not sure, and one way of doing so is to say "Do you think it might be X%?" It is likely that interviewers who were exposed to the training were more likely to suggest 10% as a possible answer.

5.2 Encouraging Guesses

Another opportunity for interviewer knowledge to affect respondents' recorded beliefs comes from the structure of our subjective belief elicitation questions. These were designed so that whenever respondents answered 50% to any risk belief questions, they triggered a follow-up question that asked whether they really thought the answer was 50%, or whether they were just unsure. If respondents said they were just unsure, they were asked for their best guess. This approach was adapted from the US Health and Retirement Survey (HRS), with the goal of reducing the use of 50% as a proxy for respondent uncertainty; see Hudomiet, Kézdi and Willis (2011) for a discussion of this technique.

These follow-ups initiated another interaction between the interviewer and respondent, creating an additional opportunity for interviewer knowledge to spill over onto survey responses. Table 4 shows our exploration of that additional interaction, in the case of per-act, unprotected transmission risks. We created indicator variables for when respondents answered 50% (column 1), changed, decreased or increased their answer after the follow-up (columns two through 4). Columns 5 and 6 show whether respondents decreased or increased their responses, for the subset of people who originally answered 50%.

Respondents in the treatment and control groups are equally likely to answer 50%, and equally likely to change their answer after the follow-up (column 2). However, respondents that were exposed to this additional interaction were significantly less likely to increase the answer

[9] Similarly, previous research has shown that subjective beliefs are affected by the exact wording of the question, i.e. framing a question as being about mortality vs. survival (Delavande et al. 2017).

after the follow up when they were interviewed by a trained interviewer, shown in columns 4 and 6. Column 6 shows that conditional on initially answering 50%, respondents exposed to informed interviewers were almost 20 percentage points less likely to increase their answer. This magnitude is large, considering that only about 30 percent of those in the control group increased their answers after the follow-up.

We interpret these results as additional evidence that interviewers who had been exposed to the information treatment influenced the responses given by communicating the follow-up question in a way that nudges or primes respondents to not increase their answers. This could be anything from a change in tone of voice or body language to the choice of words.[10] A specific possibility is that instead of asking whether the number could be more or less than 50%, they only asked if it could be less. We do not believe interviewers did this intentionally, as they knew the purpose of the intervention was to study the respondents' knowledge and behavior. Rather we believe that interviewers inadvertently nudged respondents towards lower answers. Equivalently, it could be that interviewers who had not been exposed to the information treatment nudged respondents to provide higher answers—potentially stemming from their own beliefs prior to the information treatment.

5.3 Interviewer Knowledge and Respondent Priors

If interviewer knowledge spillovers indeed operate through nudges and priming that take place during the survey interview, we would expect the effects to be smaller for respondents who are more-confident in their beliefs. Because our outcomes are measured on the baseline survey, we do not have direct measures of respondents' beliefs in the absence of the knowledge spillovers. However, some of their other characteristics may be useful proxies. Table 5 examines treatment effect heterogeneity by a range of respondent characteristics, estimated using equation 2. We observe significant heterogeneity by years of schooling and total assets.

These characteristics are correlated with one another, and thus we may be finding spurious heterogeneity by some characteristics due to omitted variable bias. Therefore we include all eight

[10] Ideally, we would have directly observed some interviews to measure the micro-processes that drove the knowledge spillovers. We did not do this for two reasons. First, the study was not designed to measure these spillovers. Second, direct participation in the survey by outsiders, especially white foreigners, can itself affect respondent behavior (Cilliers et al. 2015).

interactions in column 9, and in column 10 we add interactions with additional characteristics as well. In our preferred specification, column 10, only the interaction between treatment and years of schooling remains significant.[11] The positive coefficient for years of schooling means that the main treatment effect (the effect of having a more knowledgeable interviewer) is smaller in magnitude for those with more higher education.

To further explore this finding, we run another set of regressions with the dependent variable being beliefs about per-act unprotected HIV transmission risks, and the independent variables including the treatment interacted with seven different measures of schooling: years of schooling, having completed at least Form 1 or Form 2, and having completed exactly Form 1 through 4.[12] The regression results can be found in Online Appendix Table 9. The results suggest that the treatment effects are lowest for people who reached Form 2, which may include some of the value of 1 when the respondent has completed at least Form 2, regardless of whether the respondent continued their education beyond Form 2. Note that despite the fact that our survey respondents were all adults, most had not attended secondary school. Just 20% had finished Form 1, and only 17% had finished Form 2.

Form 2 is the point at which students in Malawi are most exposed to information on HIV transmission and sexual health.[13] NGOs in Malawi also tend to target students of this age for HIV-prevention interventions; in other African countries it is also common to target HIV-prevention campaigns at students early in secondary school (Gallant and Maticka-Tyndale 2004). The narrative to which students in Malawi are exposed in these lectures and courses is that HIV is highly contagious. This should lead to high beliefs and high certainty about those beliefs.[14]

---

[11] The standard errors in this specification may be overstated due to multicollinearity between age, years of schooling, and years sexually active. Online Appendix Table 8 shows that the condition number for the three variables is nearly 18; a figure above 10 can indicate that coefficients are unstable. However, the variance inflation factors are all well below the usual cutoff of 10.
[12] Form 1 in Malawi is the equivalent of 9th grade in the United States.
[13] HIV education was moved from other subjects into a course called life skills in the early 2000s (Chamba 2009). When this change was initially rolled out in 2001-2002, HIV was included only in the secondary school life skills curriculum. The current life skills curriculum in upper primary school (grades P5-P8) is supposed to include HIV education but there are many constraints to implementation (Chirwa and Naidoo 2014). Based on conversations with MoE officials in 2012, at that time HIV education was only done in secondary schools. An examination of the textbooks for the secondary school life skills courses revealed HIV content in all four grades, but HIV transmission risks were covered only in Form 2 (Kadyoma et al. 2012).
[14] This idea is supported by the correlation between risk beliefs and schooling for the control group: more schooling is associated with higher priors. (Online Appendix Table 10).

Other people have likely heard about the information provided to students in Form 2, for example by hearing about it from their friends. There are also other institutions that attempt to teach about HIV transmission, such as NGOs and church groups, and so other people may also have strong priors about HIV transmission risks. However, direct exposure to this information in school could lead people to be more certain about it, and thus less susceptible to nudges by interviewers.[15]

6 Preventing and Mitigating Interviewer Knowledge Effects

What should we do about the fact that interviewer knowledge spills over onto the subjective beliefs recorded on face-to-face surveys? There are a variety of ways to combat this potential spillover. First, researchers can try to alter how interviewers are recruited and trained. Recruitment matters because these spillovers can occur whenever interviewers and respondents differ in their knowledge levels. Interviewers should be recruited in a way that matches the respondent population as closely as possible, in particular in terms of education and exposure to information relevant to the survey questions. This will prevent knowledge effects, assuming that the effect we measure is driven by gaps between what interviewers know and what respondents know; in that case, eliminating differences in knowledge should correct for the issue. Controlling for interviewer fixed effects can help eliminate the effects of any remaining knowledge differences, by purging the results of any interview-specific patterns. The survey design should also be mindful of this possibility and have exact scripts for asking belief questions, to minimize selective nudges by the interviewer. Training sessions should emphasize the potential for these spillovers and coach interviewers on how to avoid them.

Second, the problem can be tackled through changes in the design of experiments when studies involve information treatments. Possible solutions include either running the baseline surveys simultaneously across the treatment and control groups or separating the information treatment from the survey data collection entirely. Each strategy has important potential drawbacks. Running simultaneous surveys across study arms creates the possibility that

[15] Another potential explanation for the fact that more-educated people respond less to the treatment is that education may make people more confident and more able to stand up to outsiders. We cannot directly test this explanation against the effect of HIV education on the strength of people's priors.

respondents will be given the wrong version of the survey and thus be unintentionally exposed to the information treatment, creating a far worse contamination problem. Running the information treatment separately from the survey, for example via village meetings, can make it difficult to prevent non-targeted people from receiving the information. If there are diagrams that are distributed as handouts, these could potentially make their way into the hands of control-group subjects. Thus incorporating information treatments into surveys is likely to *minimize* information spillovers, not exacerbate them, but at the cost of yielding potentially-biased measurements of respondent beliefs. If accurate measures of respondent beliefs are not a main goal of the study, this may be an acceptable risk. For example, if the goal of an experiment is to see how much an information treatment shifts behaviors, then mismeasured beliefs are not a problem, even if they affect only one of the study arms. Even if looking at treatment effects on risk beliefs is an important goal of the study, interviewer knowledge contamination is only a problem if it interacts with the actual treatment. Apart from information experiments, knowledge spillovers are likely to occur simply because interviewers differ in their knowledge and beliefs. Providing a basic level of knowledge about important survey topics could help them do their jobs better and lead to better-quality data.

A third solution to this issue is to collect subjective expectations in a way that avoids any direct interaction with interviewers, such as by using computer-assisted self-interviewing (CASI). This would eliminate any possibility of interviewer knowledge spilling over onto respondents. Research on CASI has shown it can be effective in low-literacy settings (van de Wijgert et al. 2000, Hahn et al. 2003). There are important limitations, however: participants may not be able to clarify questions (NIMH 2007), computers may be received with suspicion in certain settings (Mensch et al. 2003, Hewett et al. 2004), and bystander presence might affect results and should be recorded or controlled (Aquilino et al 2000). Potdar and Koenig (2005) argue that CASI will not yield more-honest answers if people are not comfortable using computers. In low-literacy settings, audio computer-assisted self-interviewing (ACASI) may work well; Rumakom et al (2005) shows that it outperforms self-administered questionnaires. However, Soler-Hampejsek et al. (2013) test ACASI for collecting sexual activity data in Malawi and find that it still leads to high rates of inconsistency in responses. Similarly, Mensch et al. (2008) show that face-to-face interviews generate lower rates of inconsistencies in responses than ACASI, and stronger correlations between reported sexual behavior and biomarkers for HIV infection. To improve the quality of subjective expectation data in developing countries, more work on adapting CASI and ACASI to overcome these limitations is needed. One promising approach is to use tablets to have respondents play simple games that convey information; Tjernström et al. (2019) show that this approach is successful in a population of Kenyan farmers.

7 Conclusion

Leveraging a randomized experiment that used interviewers to measure subjective HIV transmission risks and provide information to treatment group participants, we find that interviewer knowledge affects the recorded values of survey respondents' subjective beliefs. This information spillover happens not only for the information directly given to the interviewers, but for other related risks.

We identify several channels through which these effects happen. They are evident at various points in the survey, including the follow up questions triggered by respondents answering 50%. This result suggests that additional interactions between interviewers and respondents present the potential for more spillovers. Our evidence suggests that interviewer effects work via priming or nudging rather than interviewers directly revealing information.

We find that interviewer effects are weaker for more-educated people, possibly because those respondents received information about HIV transmission directly at school and are more certain about their prior beliefs than those who heard information second hand, even if the level of those beliefs is not different across education levels.

Our results have important implications for demographers as well as other social scientists who study subjective expectations themselves or phenomena that are driven by people's subjective beliefs. Subjective expectations have proven to be useful tools for understanding and forecasting the main demographic processes of fertility, mortality, and migration, but these uses rely on being able to measure them correctly. Researchers need to be aware of the possibility that interviewer knowledge will spill over onto respondents' recorded beliefs, which could have substantive effects on results that use those beliefs. While our findings are for HIV risk perceptions, interviewer knowledge effects could occur for any subjective expectation where the interviewer knows more than the respondent—including other diseases such as Ebola, or COVID-19, and also other domains where subjective expectations play a role like conception probabilities, mortality rates, and the returns to migration. Our results are for a setting where interviewers were evaluated based on recording responses correctly; they may not generalize to settings where interviewers have some interest in recording specific responses.

We suggest methods for correcting this problem at several points during the course of a research project. These include mindful recruiting of interviewers to match the knowledge levels of the respondent population, emphasizing the potential for spillovers in training, and designing the experiment in such a way that interviewers survey both study arms while having the same information set. The most promising way to avoid interviewer knowledge effects is to collect data via CASI or ACASI, to reduce the scope of interaction between respondent and interviewer, but both these methods have issues with data quality. Interviewer knowledge effects are therefore likely to remain an issue for measuring subjective beliefs in developing-country settings for the foreseeable future.

Acknowledgments

We thank Adeline Delavande, Audrey Dorélien, Jennifer Johnson-Hanks, Maxwell Mkondiwa, Stacy Pancratz, Rebecca Thornton, Emilia Tjernström, Jenny Trinitapoli, Susan Watkins, Rob Warren, Bob Willis, Elizabeth Wrigley-Field, and four anonymous reviewers for their 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). All errors and omissions are our own.

Conflict of Interest Statement

The authors declare that there are no potential conflicts of interest.

Data Availability

The data used in this paper is available online via the Harvard Dataverse: Kerwin, Jason, and Natalia Ordaz Reynoso. 2020. Replication Data for: You Know What I Know: Interviewer Knowledge Effects in Subjective Expectation Elicitation. Harvard Dataverse. https://doi.org/10.7910/DVN/8TSPEJ

References

Adida, C. L., Ferree, K. E., Posner, D. N., & Robinson, A. L. (2016). Who's Asking? Interviewer Coethnicity Effects in African Survey Data. *Comparative Political Studies*, *49*(12), 1630–1660. https://doi.org/10.1177/0010414016633487

Aguila, E., Borges, A., Castillejos, C. M., Pierson, A., & Weidmer, B. A. (2014). *Mortality Expectations of Older Mexicans: Development and Testing of Survey Measures* [Product Page]. https://www.rand.org/pubs/technical_reports/TR1288z6.html

Alhaj, M., Amberbir, A., Singogo, E., Banda, V., van Lettow, M., Matengeni, A., Kawalazira, G., Theu, J., Jagriti, M. R., Chan, A. K., & van Oosterhout, J. J. (2019). Retention on antiretroviral therapy during Universal Test and Treat implementation in Zomba district, Malawi: A retrospective cohort study. *Journal of the International AIDS Society*, *22*(2). https://doi.org/10.1002/jia2.25239

Anderson, B. A., Silver, B. D., & Abramson, P. R. (1988). The Effects of the Race of the Interviewer on Race-related Attitudes of Black Respondents in SRC/CPS National Election Studies. *Public Opinion Quarterly*, *52*(3), 289–324. https://doi.org/10.1086/269108

Andreenkova, A. V. (2018). How to Choose Interview Language in Different Countries. In T. P. Johnson, B.-E. Pennell, I. A. Stoop, & B. Dorer (Eds.), *Advances in comparative survey methods: Multinational, multiregional, and multicultural contexts (3MC)*. John Wiley & Sons.

Anglewicz, P., & Chintsanya, J. (2011). Disclosure of HIV status between spouses in rural Malawi. *AIDS Care*, *23*(8), 998–1005. https://doi.org/10.1080/09540121.2010.542130

Anglewicz, P., & Clark, S. (2013). The effect of marriage and HIV risks on condom use acceptability in rural Malawi. *Social Science & Medicine*, *97*, 29–40.

https://doi.org/10.1016/j.socscimed.2013.06.024

Anglewicz, P., & Kohler, H.-P. (2009). Overestimating HIV infection: The construction and accuracy of subjective probabilities of HIV infection in rural Malawi. *Demographic Research*, *20*(6), 65–96. https://doi.org/10.4054/DemRes.2009.20.6

Anglewicz, P., & Reniers, G. (2014). HIV Status, Gender, and Marriage Dynamics among Adults in Rural Malawi. *Studies in Family Planning*, *45*(4), 415–428. https://doi.org/10.1111/j.1728-4465.2014.00005.x

Angotti, N., Dionne, K. Y., & Gaydosh, L. (2011). An offer you can't refuse? Provider-initiated HIV testing in antenatal clinics in rural Malawi. *Health Policy and Planning*, *26*(4), 307–315.

Aquilino, W. S., Wright, D. L., & Supple, A. J. (2000). Response effects due to bystander presence in CASI and paper-and-pencil surveys of drug use and alcohol use. *Substance Use & Misuse*, *35*(6–8), 845–867.

Attanasio, O., Meghir, C., & Vera-Hernández, M. (2005). *Elicitation, validation, and use of probability distributions of future income in developing countries* [Working Paper].

Attanasio, O. P. (2009). Expectations and Perceptions in Developing Countries: Their Measurement and Their Use. *American Economic Review*, *99*(2), 87–92. https://doi.org/10.1257/aer.99.2.87

Becker, S., Feyisetan, K., & Makinwa-Adebusoye, P. (1995, July 1). *The effect of the sex of interviewers on the quality of data in a Nigerian family planning questionnaire*. Studies in Family Planning. https://link.galegroup.com/apps/doc/A17474999/AONE?sid=lms

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

Bennear, L., Tarozzi, A., Pfaff, A., Balasubramanya, S., Matin Ahmed, K., & van Geen, A. (2013). Impact of a randomized controlled trial in arsenic risk communication on household water-source choices in Bangladesh. *Journal of Environmental Economics and Management*, *65*(2), 225–240. https://doi.org/10.1016/j.jeem.2012.07.006

Bignami-Van Assche, S., Chao, L. W., Anglewicz, P., Chilongozi, D., & Bula, A. (2007). The validity of self-reported likelihood of HIV infection among the general population in rural Malawi. *Sexually Transmitted Infections*, *83*(1), 35–40.

Blaydes, L., & Gillum, R. M. (2013). Religiosity-of-Interviewer Effects: Assessing the Impact of Veiled Enumerators on Survey Response in Egypt. *Politics and Religion*, *6*(3), 459–482. https://doi.org/10.1017/S1755048312000557

Burström, B., & Fredlund, P. (2001). Self rated health: Is it as good a predictor of subsequent mortality among adults in lower as well as in higher social classes? *Journal of Epidemiology & Community Health*, *55*(11), 836–840. https://doi.org/10.1136/jech.55.11.836

Calonico, S., Cattaneo, M. D., Farrell, M. H., & Titiunik, R. (2017). Rdrobust: Software for Regression-discontinuity Designs. *The Stata Journal: Promoting Communications on Statistics and Stata*, *17*(2), 372–404. https://doi.org/10.1177/1536867X1701700208

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica*, *82*(6), 2295–2326. https://doi.org/10.3982/ECTA11757

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2015). Optimal Data-Driven Regression Discontinuity Plots. *Journal of the American Statistical Association*, *110*(512), 1753–1769. https://doi.org/10.1080/01621459.2015.1017578

Chamba, M. V. M. (2009). *Evaluation Study on Implementation of Life Skills Programme in Public*

Secondary Schools in Malawi: South Eastern Educational Division (SEED). University of Malawi College of Medicine.

Chapman, G. B., & Johnson, E. J. (2002). Incorporating the Irrelevant: Anchors in Judgments of Belief and Value. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and Biases* (1st ed., pp. 120–138). Cambridge University Press. https://doi.org/10.1017/CBO9780511808098.008

Chinkhumba, J., Godlonton, S., & Thornton, R. (2014). Demand for medical male circumcision. *American Economic Journal: Applied Economics*, *6*(2), 152–177.

Chirwa, G., & Naidoo, D. (2014). *Structural and social constraints in the teaching of Life Skills for HIV/AIDS prevention in Malawi primary schools*. 19.

Chuang, Y., & Schechter, L. (2015). Stability of experimental and survey measures of risk, time, and social preferences: A review and some new results. *Journal of Development Economics*, *117*, 151–170. https://doi.org/10.1016/j.jdeveco.2015.07.008

Cilliers, J., Dube, O., & Siddiqi, B. (2015). The white-man effect: How foreigner presence affects behavior in experiments. *Journal of Economic Behavior & Organization*, *118*, 397–414. https://doi.org/10.1016/j.jebo.2015.03.015

Cotter, P. R., Cohen, J., & Coulter, P. B. (1982). Race-of-Interviewer Effects in Telephone Interviews. *Public Opinion Quarterly*, *46*(2), 278–284. https://doi.org/10.1086/268719

Davis, D. W. (1997). Nonrandom Measurement Error and Race of Interviewer Effects Among African Americans. *The Public Opinion Quarterly*, *61*(1), 183–207. JSTOR.

de Mel, S., McKenzie, D., & Woodruff, C. (2008). Returns to Capital in Microenterprises: Evidence from a Field Experiment. *The Quarterly Journal of Economics*, *123*(4), 1329–1372. https://doi.org/10.1162/qjec.2008.123.4.1329 de Quidt, J., Haushofer, J., & Roth, C. (2018). Measuring and Bounding Experimenter Demand. *American Economic Review*, *108*(11), 3266–3302. https://doi.org/10.1257/aer.20171330

Delavande, A. (2008). Pill, patch, or shot? Subjective expectations and birth control choice. *International Economic Review*, *49*(3), 999–1042.

Delavande, A. (2014). Probabilistic Expectations in Developing Countries. *Annual Review of Economics*, *6*(1), 1–20. https://doi.org/10.1146/annurev-economics-072413-105148

Delavande, A., Giné, X., & McKenzie, D. (2011a). Eliciting probabilistic expectations with visual aids in developing countries: How sensitive are answers to variations in elicitation design? *Journal of Applied Econometrics*, *26*(3), 479–497. https://doi.org/10.1002/jae.1233

Delavande, A., Giné, X., & McKenzie, D. (2011b). Measuring subjective expectations in developing countries: A critical review and new evidence. *Journal of Development Economics*, *94*(2), 151–163. https://doi.org/10.1016/j.jdeveco.2010.01.008

Delavande, A., & Kohler, H.-P. (2009). Subjective expectations in the context of HIV/AIDS in Malawi. *Demographic Research*, *20*, 817–874. https://doi.org/10.4054/DemRes.2009.20.31

Delavande, A., & Kohler, H.-P. (2012). The Impact of HIV Testing on Subjective Expectations and Risky Behavior in Malawi. *Demography*, *49*(3), 1011–1036. https://doi.org/10.1007/s13524-012-0119-7

Delavande, A., & Kohler, H.-P. (2016). HIV/AIDS-related Expectations and Risky Sexual Behaviour in Malawi. *The Review of Economic Studies*, *83*(1), 118–164. https://doi.org/10.1093/restud/rdv028

Delavande, A., Lee, J., & Menon, S. (2017). Eliciting Survival Expectations of the Elderly in Low-Income Countries: Evidence From India. *Demography*, *54*(2), 673–699. https://doi.org/10.1007/s13524-017-0560-8

Delavande, A., & Rohwedder, S. (2011). Differential Survival in Europe and the United States: Estimates Based on Subjective Probabilities of Survival. *Demography*, *48*(4), 1377–1400. https://doi.org/10.1007/s13524-011-0066-8

Di Maio, M., & Fiala, N. (2019). Be Wary of Those Who Ask: A Randomized Experiment on the Size and Determinants of the Enumerator Effect. *The World Bank Economic Review*, *in press*. https://doi.org/10.1093/wber/lhy024

Dionne, K. Y. (2014). The politics of local research production: Surveying in a context of ethnic competition. *Politics, Groups, and Identities*, *2*(3), 459–480. https://doi.org/10.1080/21565503.2014.930691

Dormont, B., Samson, A.-L., Fleurbaey, M., Luchini, S., & Schokkaert, E. (2018). Individual Uncertainty About Longevity. *Demography*, *55*(5), 1829–1854. https://doi.org/10.1007/s13524-018-0713-4

Elder, T. E. (2013). The Predictive Validity of Subjective Mortality Expectations: Evidence From the Health and Retirement Study. *Demography*, *50*(2), 569–589. https://doi.org/10.1007/s13524-012-0164-2

Fedor, T. M., Kohler, H.-P., & Behrman, J. R. (2015). The Impact of Married Individuals Learning HIV Status in Malawi: Divorce, Number of Sexual Partners, and Condom Use With Spouses. *Demography*, *52*(1), 259–280. https://doi.org/10.1007/s13524-014-0364-z

Finkel, S. E., Guterbock, T. M., & Borg, M. J. (1991). RACE-OF-INTERVIEWER EFFECTS IN A PREELECTION POLL VIRGINIA 1989. *Public Opinion Quarterly*, *55*(3), 313–330. https://doi.org/10.1086/269264

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

Gallant, M., & Maticka-Tyndale, E. (2004). School-based HIV prevention programmes for African youth. *Social Science & Medicine*, *58*(7), 1337–1351. https://doi.org/10.1016/S0277-9536(03)00331-9

Godlonton, S., Munthali, A., & Thornton, R. (2015). Responding to Risk: Circumcision, Information, and HIV Prevention. *Review of Economics and Statistics*, in press.

Grant, M. J. (2008). Children's school participation and HIV/AIDS in rural Malawi: The role of parental knowledge and perceptions. *Demographic Research*, *19*, 1603–1634. https://doi.org/10.4054/DemRes.2008.19.45

Grant, M. J., & Soler-Hampejsek, E. (2014). HIV Risk Perceptions, the Transition to Marriage, and Divorce in Southern Malawi. *Studies in Family Planning*, *45*(3), 315–337. https://doi.org/10.1111/j.1728-4465.2014.00394.x

Hahn, E. A., Cellal, D., Dobrez, D. G., Shiomoto, G., Taylor, S. G., Galvez, A. G., Diaz, P., Valenzuela, V., Chiang, H. L., Khan, S., Hudgens, S. A., & Du, H. (2003). Quality of life assessment for low literacy Latinos: A new multimedia program for self-administration. *The Journal of Oncology Management : The Official Journal of the American College of Oncology Administrators*, *12*(5), 9–12.

Helleringer, S., & Kohler, H.-P. (2005). Social networks, perceptions of risk, and changing attitudes towards HIV/AIDS: New evidence from a longitudinal study using fixed-effects analysis. *Population Studies*, *59*(3), 265–282. https://doi.org/10.1080/00324720500212230

Hewett, P. C., Erulkar, A. S., Mensch, B. S., & Kenya, C. (2004). The feasibility of computer-assisted survey interviewing in Africa: Experience from two rural districts in Kenya. *Social Science Computer Review*.

Hudomiet, P., Kézdi, G., & Willis, R. J. (2011). Stock Market Crash and Expectations of American

Households. *Journal of Applied Econometrics*, *26*(3), 393–415. https://doi.org/10.1002/jae.1226

Hurd, M. D., & McGarry, K. (1995). Evaluation of the Subjective Probabilities of Survival in the Health and Retirement Study. *The Journal of Human Resources*, *30*, S268–S292. JSTOR. https://doi.org/10.2307/146285

Idler, E. L., & Benyamini, Y. (1997). Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies. *Journal of Health and Social Behavior*, *38*(1), 21–37. JSTOR. https://doi.org/10.2307/2955359

Imbens, G. W., & Rubin, D. B. (2015, April). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge Core. https://doi.org/10.1017/CBO9781139025751

Johnson, T. P., Pennell, B.-E., Stoop, I. A., & Dorer, B. (Eds.). (2018). *Advances in comparative survey methods: Multinational, multiregional, and multicultural contexts (3MC)*. John Wiley & Sons.

Johnson-Hanks, J. A. (2017). Vital Conjunctures Revisited. In P. Kreager & A. Bochow (Eds.), *Fertility, Conjuncture, Difference: Anthropological Approaches to the Heterogeneity of Modern Fertility Declines* (1st ed., pp. 324–338). Berghahn Books. https://doi.org/10.2307/j.ctvw04c56

Kadyoma, F., Mbugua, M., Ndung'u, A., & Omae, E. (2012). *Malawi Junior Secondary Life Skills: Form 2*. Longhorn Publishers.

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

Kaler, A., & Watkins, S. (2010). Asking God about the date you will die: HIV testing as a zone of uncertainty in rural Malawi. *Demographic Research*, *23*.

Kaler, A., Watkins, S. C., & Angotti, N. (2015). Making meaning in the time of AIDS: Longitudinal narratives from the Malawi Journals Project. *African Journal of AIDS Research*, *14*(4), 303–314. https://doi.org/10.2989/16085906.2015.1084342

Kerwin, J. (2018). *Scared Straight or Scared to Death? The Effect of Risk Beliefs on Risky Behaviors* [Working Paper]. University of Minnesota.

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

Keskin, P., Shastry, G. K., & Willis, H. (2016). Water Quality Awareness and Breastfeeding: Evidence of Health Behavior Change in Bangladesh. *Review of Economics and Statistics*. https://doi.org/10.1162/REST_a_00626

Kohler, H.-P., Behrman, J. R., & Watkins, S. C. (2007). Social networks and HIV/AIDS risk perceptions. *Demography*, *44*(1), 1–33. https://doi.org/10.1353/dem.2007.0006

Kremer, M., & Miguel, E. (2007). The Illusion of Sustainability. *The Quarterly Journal of Economics*, *122*(3), 1007–1065. https://doi.org/10.1162/qjec.122.3.1007

Kuziemko, I., Pan, J., Shen, J., & Washington, E. (2018). *The Mommy Effect: Do Women Anticipate the Employment Effects of Motherhood?* (Working Paper No. 24740; Working Paper Series). National Bureau of Economic Research. https://doi.org/10.3386/w24740

León, G., & Miguel, E. (2017). Risky Transportation Choices and the Value of a Statistical Life. *American Economic Journal: Applied Economics*, *9*(1), 202–228. https://doi.org/10.1257/app.20160140

Lillard, L. A., & Willis, R. J. (2001). *Cognition and Wealth: The Importance of Probabilistic*

Thinking (SSRN Scholarly Paper ID 1084559). Social Science Research Network. https://papers.ssrn.com/abstract=1084559

Luke, N., Clark, S., & Zulu, E. M. (2011). The Relationship History Calendar: Improving the scope and quality of data on youth sexual behavior. *Demography*, *48*(3), 1151–1176.

Mac Dougall, K., Beyene, Y., & Nachtigall, R. D. (2013). Age shock: Misperceptions of the impact of age on fertility before and after IVF in women who conceived after age 40. *Human Reproduction*, *28*(2), 350–356. https://doi.org/10.1093/humrep/des409

Mahony, M. (2011). *Fertility Knowledge Among US Women Aged 25–35: Insights from a New Generation*. Medaxial. https://www.npr.org/assets/news/2011/11/FertilityWhitePaper_Final.pdf

Maluwa-Banda, D. (2004). Gender Sensitive Educational Policy and Practice: The Case of Malawi. *PROSPECTS*, *34*(1), 71–84. https://doi.org/10.1023/B:PROS.0000026680.92752.3b

Manski, C. F. (2004). Measuring Expectations. *Econometrica*, *72*(5), 1329–1376. https://doi.org/10.1111/j.1468-0262.2004.00537.x

McCombie, S. C., & Anarfi, J. K. (2002). The influences of sex of interviewer on the results of an AIDS survey in Ghana. *Human Organization; Oklahoma City*, *61*(1), 51–57.

McKenzie, D., Gibson, J., & Stillman, S. (2013). A land of milk and honey with streets paved with gold: Do emigrants have over-optimistic expectations about incomes abroad? *Journal of Development Economics*, *102*, 116–127. https://doi.org/10.1016/j.jdeveco.2012.01.001

McKenzie, D. J., Gibson, J., & Stillman, S. (2006). *How Important is Selection? Experimental vs. Non-Experimental Measures of the Income Gains from Migration* (SSRN Scholarly Paper ID 898591). Social Science Research Network. https://papers.ssrn.com/abstract=898591

Mensch, B. S., Hewett, P. C., & Erulkar, A. S. (2003). The Reporting of Sensitive Behavior by Adolescents: A Methodological Experiment in Kenya. *Demography*, *40*(2), 247–268. JSTOR. https://doi.org/10.2307/3180800

Mensch, B. S., Hewett, P. C., Gregory, R., & Helleringer, S. (2008). Sexual Behavior and STI/HIV Status Among Adolescents in Rural Malawi: An Evaluation of the Effect of Interview Mode on Reporting. *Studies in Family Planning*, *39*(4), 321–334. https://doi.org/10.1111/j.1728-4465.2008.00178.x

Minnesota Population Center. (2019). *Integrated Public Use Microdata Series, International: Version 7.2 [dataset]*. IPUMS.

Mueller, M. W., Hicks, J. H., Johnson-Hanks, J., & Miguel, E. (2019). *The Illusion of Stable Preferences over Major Life Decisions* (Working Paper No. 25844). National Bureau of Economic Research. https://doi.org/10.3386/w25844

Munthali, A. C., & Zulu, E. M. (2007). The Timing and role of Initiation Rites in Preparing Young People for Adolescence and Responsible Sexual and Reproductive Behaviour in Malawi. *African Journal of Reproductive Health*, *11*(3), 150.

Mussweiler, T., Strack, F., & Pfeiffer, T. (2000). Overcoming the Inevitable Anchoring Effect: Considering the Opposite Compensates for Selective Accessibility. *Personality and Social Psychology Bulletin*, *26*(9), 1142–1150. https://doi.org/10.1177/01461672002611010

National Statistical Office - NSO/Malawi and ICF. (2017). *Malawi Demographic and Health Survey*. NSO and ICF. https://dhsprogram.com/pubs/pdf/FR319/FR319.pdf

Newell, B. R., & Shanks, D. R. (2014). Prime Numbers: Anchoring and its Implications for Theories of Behavior Priming. *Social Cognition*, *32*(Supplement), 88–108. https://doi.org/10.1521/soco.2014.32.supp.88

NIMH Collaborative HIV/STD Prevention Trial Group. (2007). The feasibility of audio computer-assisted self-interviewing in international settings: *AIDS*, *21*(Suppl 2), S49–S58. https://doi.org/10.1097/01.aids.0000266457.11020.f0

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, *17*(11), 776–783. https://doi.org/10.1037/h0043424

Perozek, M. (2008). Using subjective expectations to forecast longevity: Do survey respondents know something we don't know? *Demography*, *45*(1), 95–113. https://doi.org/10.1353/dem.2008.0010

Potdar, R., & Koenig, M. A. (2005). Does Audio-CASI Improve Reports of Risky Behavior? Evidence from a Randomized Field Trial Among Young Urban Men in India. *Studies in Family Planning*, *36*(2), 107–116. https://doi.org/10.1111/j.1728-4465.2005.00048.x

Reese, S. D., Danielson, W. A., Shoemaker, P. J., Chang, T.-K., & Hsu, H.-L. (1986). Ethnicity-of-Interviewer Effects Among Mexican-Americans and Anglos. *Public Opinion Quarterly*, *50*(4), 563–572. https://doi.org/10.1086/269004

Rumakom, P., Guest, P., Chinvarasopak, W., Utarmat, W., & Sontanakanit, J. (2005). Obtaining accurate responses to sensitive questions among Thai students: A comparison of two data collection techniques. In S. Jejeebhoy, I. Shah, & S. Thapa (Eds.), *Sex Without Consent* (pp. 318–332). Zed Books.

Shapira, G. (2017). How Subjective Beliefs about HIV Infection Affect Life-Cycle Fertility: Evidence from Rural Malawi. *Journal of Human Resources*, *52*(3), 680–718. https://doi.org/10.3368/jhr.52.3.0314-6263R1

Shrestha, M. (2020). Get Rich or Die Tryin': Perceived Earnings, Perceived Mortality Rates, and

Migration Decisions of Potential Work Migrants from Nepal. *The World Bank Economic Review*, *34*(1), 1–27. https://doi.org/10.1093/wber/lhz023

Smith, K. P., & Watkins, S. C. (2005). Perceptions of risk and strategies for prevention: Responses to HIV/AIDS in rural Malawi. *Social Science & Medicine*, *60*(3), 649–660. https://doi.org/10.1016/j.socscimed.2004.06.009

Soler-Hampejsek, E., Grant, M. J., Mensch, B. S., Hewett, P. C., & Rankin, J. (2013). The Effect of School Status and Academic Skills on the Reporting of Premarital Sexual Behavior: Evidence From a Longitudinal Study in Rural Malawi. *Journal of Adolescent Health*, *53*(2), 228–234. https://doi.org/10.1016/j.jadohealth.2013.03.008

Thornton, R. L. (2008). The Demand for, and Impact of, Learning HIV Status. *American Economic Review*, *98*(5), 1829–1863.

Tjernström, E., Lybbert, T. J., Hernández, R. F., & Correa, J. S. (2019). *Learning by (virtually) doing: Experimentation and belief updating in smallholder agriculture* [Working Paper].

Trinitapoli, J. (2009). Religious teachings and influences on the ABCs of HIV prevention in Malawi. *Social Science & Medicine*, *69*(2), 199–209. https://doi.org/10.1016/j.socscimed.2009.04.018

Trinitapoli, J., & Regnerus, M. D. (2006). Religion and HIV Risk Behaviors Among Married Men: Initial Results from a Study in Rural Sub-Saharan Africa. *Journal for the Scientific Study of Religion*, *45*(4), 505–528. https://doi.org/10.1111/j.1468-5906.2006.00325.x

Trinitapoli, J., & Yeatman, S. (2011). Uncertainty and Fertility in a Generalized AIDS Epidemic. *American Sociological Review*, *76*(6), 935–954. https://doi.org/10.1177/0003122411427672

Trinitapoli, J., & Yeatman, S. (2018). The Flexibility of Fertility Preferences in a Context of Uncertainty: Flexibility of Fertility Preferences in a Context of Uncertainty. *Population and*

Development Review, *44*(1), 87–116. https://doi.org/10.1111/padr.12114

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

van de Wijgert, J., Padian, N., Shiboski, S., & Turner, C. (2000). Is audio computer-assisted self-interviewing a feasible method of surveying in Zimbabwe? *International Journal of Epidemiology*, *29*(5), 885–890.

Watkins, S. C. (2011). Back to Basics: Gender, Social Norms, and the AIDS Epidemic in Sub-Saharan Africa. In D. E. Sahn (Ed.), *The Socioeconomic Dimensions of HIV/AIDS in Africa: Challenges, Opportunities and Misconceptions* (pp. 134–162). Cornell University Press.

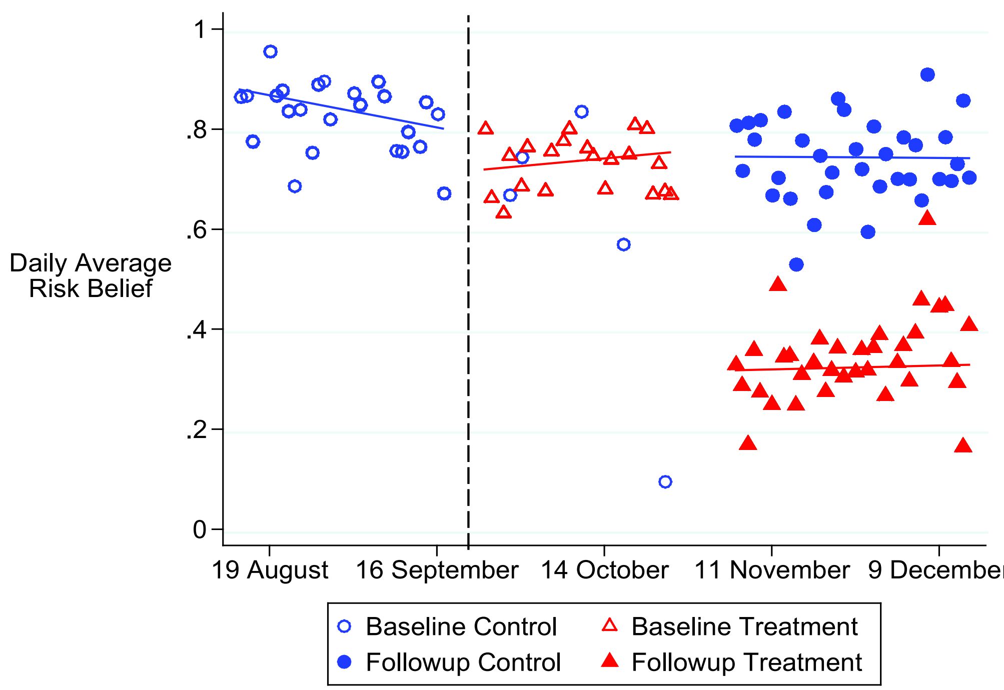
Wawer, M. J., Gray, R. H., Sewankambo, N. K., Serwadda, D., Li, X., Laeyendecker, O., Kiwanuka, N., Kigozi, G., Kiddugavu, M., Lutalo, T., Nalugoda, F., Wabwire-Mangen, F., Meehan, M. P., & Quinn, T. C. (2005). Rates of HIV-1 Transmission per Coital Act, by Stage of HIV-1 Infection, in Rakai, Uganda. *The Journal of Infectious Diseases*, *191*(9), 1403–1409. https://doi.org/10.1086/429411

Weinreb, A. A. (2006). The limitations of stranger-interviewers in rural Kenya. *American Sociological Review*, *71*(6), 1014–1039.

Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, *13*(1), 75–98. https://doi.org/10.1007/s10683-009-9230-z

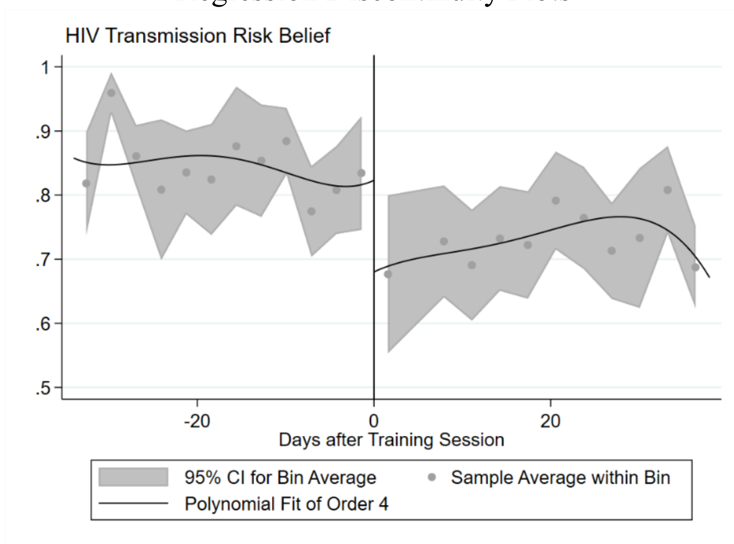
Figure 1

Measured Risk Beliefs over Time, by Study Arm

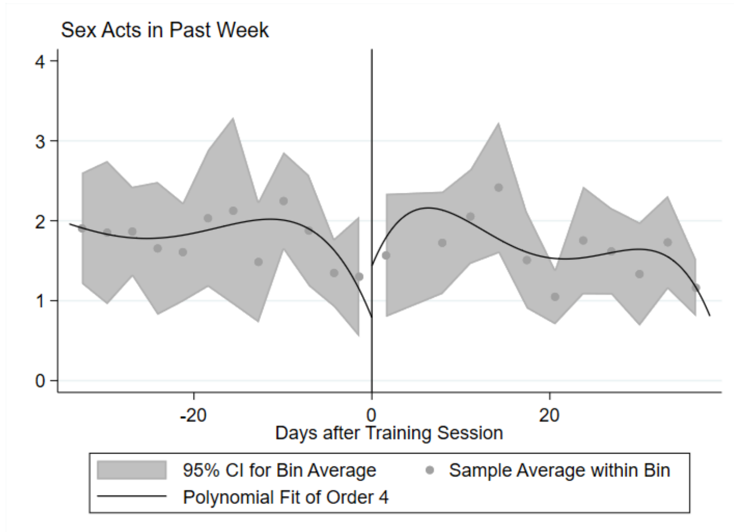


Notes: Sample includes 1,292 sexually-active adults who were successfully interviewed at both baseline and endline. Risk beliefs are the perceived probability of contracting HIV from a single unproteted sex act with an infected partner. Each point represents the mean value of the risk beliefs for a given day; baseline control beliefs are hollow circles, endline control beliefs are solid circles, baseline treatment beliefs are hollow triangles, and endline treatment beliefs are solid triangles. The lines are linear fits of beliefs on date for a given date range and study arm. The dashed vertical line indicates the date of the training sessions when the survey interviewers were trained to provide the information treatment about HIV transmission risks.

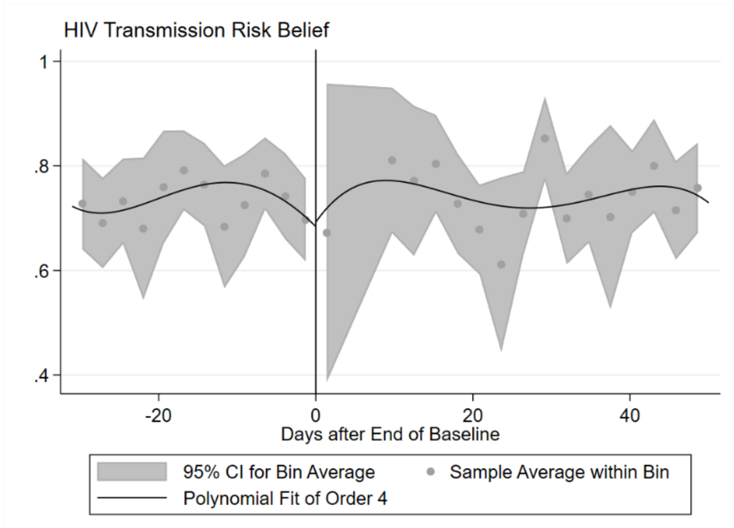
Figure 2
Regression Discontinuity Plots



Panel A: HIV Transmission Risk Belief, Discontinuity at Training Session



Panel B: Sex Acts in Past Week, Discontinuity at Training Session



Panel C: HIV Transmission Risk Belief, Discontinuity at Training Session (Control Group Only)

Notes: Sample is 1,292 sexually-active adults who were successfully interviewed at both baseline and endline. HIV transmission risk belief is the perceived probability of contracting HIV from a single unprotected sex act with an infected partner. Sex acts in past week is measured using a seven-day retrospective sex "diary."

Table 1

Effects of Interviewer Knowledge on Reported Risk Beliefs

	(1)	(2)	(3)	(4)
	Outcome: HIV transmission risk belief			
	Per Act, Unprotected	Per Year, Unprotected	Per Act, With Condom	Per Year, With Condom
Treatment	-0.0745***	-0.0352***	-0.0366***	-0.0850***
	(0.0191)	(0.0128)	(0.0128)	(0.0157)
Control-group mean	0.827	0.927	0.123	0.236
Control-group SD	0.264	0.169	0.218	0.279
Observations	1,282	1,277	1,284	1,277

Notes: All regresions control for stratification cell and interviewer fixed effects, as well as the alternate sexual behavior index from Appendix Table 1. Sample includes 1,292 sexually-active adults who were successfully interviewed at both baseline and endline. Heteroskedasticity-robust standard errors, clustered by village, in parentheses: * p<0.1; ** p<0.05; *** p<0.01.

Table 2

Regression Discontinuity Estimates

	HIV Transmission Risk Belief	Sex Acts in Past Week (from Sex "Diary")	1(Respondent is Male)	Age
Panel A: Discontinuity at Training Session				
RD Estimate	-0.148**	0.476	0.041	-0.876
Standard Error	(0.069)	(0.362)	(0.085)	(1.926)
Conventional p-value	[0.033]	[0.189]	[0.632]	[0.649]
Robust p-value	{0.034}	{0.155}	{0.698}	{0.400}
Sample:	1,289	1,292	1,292	1,292
Control-group Baseline Surveys	X	X	X	X
Treatment-group Baseline Surveys	X	X	X	X
Control-group Endline Surveys				
Treatment-group Endline Surveys				
Observations				
Panel B: Discontinuity at Training Session, Control Group Only				
RD Estimate	-0.019	0.318	-0.004	-1.379
Standard Error	(0.144)	(0.346)	(0.170)	(2.227)
Conventional p-value	[0.896]	[0.358]	[0.981]	[0.536]
Robust p-value	{0.776}	{0.173}	{0.748}	{0.386}
Sample:				
Control-group Baseline Surveys				
Treatment-group Baseline Surveys	X	X	X	X
Control-group Endline Surveys	X	X	X	X
Treatment-group Endline Surveys				
Observations	1,308	1316	1316	1316

Notes: Sample includes 1,376 sexually-active adults who were successfully interviewed at both baseline and followup. Regression discontinuity estimates generated first-degree polynomials and automatic bandwidth selection following Calonico et al. (2014). Heteroskedasticity-robust standard errors, clustered by village, in parentheses: * p<0.1; ** p<0.05; *** p<0.01.

Table 3

Treatment Effects on Answering Exactly 10%

	(1)	(2)	(3)	(4)
	Outcome: Recorded belief is exactly 10%			
	Per Act, Unprotected	Per Year, Unprotected	Per Act, With Condom	Per Year, With Condom
Treatment	0.0427***	0.00665	-0.0507***	-0.0201
	(0.0117)	(0.00601)	(0.0189)	(0.0171)
Control-group mean	0.022	0.005	0.100	0.100
Control-group SD	0.146	0.068	0.300	0.300
Observations	1,282	1,277	1,284	1,277

Notes: All regresions control for stratification cell and interviewer fixed effects, as well as the alternate sexual behavior index from Appendix Table 1. Sample includes 1,292 sexually-active adults who were successfully interviewed at both baseline and endline. Heteroskedasticity-robust standard errors, clustered by village, in parentheses: * p<0.1; ** p<0.05; *** p<0.01.

Table 4

Priming by Interviewers when Initial Answer was 50%

	Answer = 50%	Changed Answer	Decreased Answer	Increased Answer	Decreased Answer	Increased Answer
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.0161	-0.0126	0.000651	-0.0133	-0.00832	-0.115*
	(0.0223)	(0.0122)	(0.00704)	(0.00872)	(0.0472)	(0.0630)
Conditional on Initially Answering 50%	N	N	N	N	Y	Y
Control-group mean	0.115	0.042	0.008	0.034	0.0676	0.297
Control-group SD	0.319	0.201	0.088	0.182	0.253	0.460
Observations	1,285	1,285	1,285	1,285	159	159

Notes: All regresions control for stratification cell and interviewer fixed effects, as well as the alternate sexual behavior index from Appendix Table 1. Sample includes 1,292 sexually-active adults who were successfully interviewed at both baseline and endline. Heteroskedasticity-robust standard errors, clustered by village, in parentheses: * p<0.1; ** p<0.05; *** p<0.01.

Table 5

Heterogeneity in Treatment Effects

	Outcome: HIV transmission risk belief (per-act, unprotected)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment	-0.0753***	-0.0745***	-0.0756***	-0.0769***	-0.0770***	-0.0737***	-0.0746***	-0.0804***	-0.0891***	-0.106***
	(0.0191)	(0.0191)	(0.0191)	(0.0194)	(0.0200)	(0.0191)	(0.0191)	(0.0194)	(0.0210)	(0.0251)
T × (Age)	0.000458								-0.00230	-0.000713
	(0.00177)								(0.00469)	(0.00523)
T × (Male)		-0.0115							-0.0558	-0.0657*
		(0.0307)							(0.0338)	(0.0383)
T × (Years of Schooling)			0.0112***						0.00768	0.0115**
			(0.00397)						(0.00501)	(0.00557)
T × (Years Sexually Active)				0.000258					0.00422	0.00269
				(0.00175)					(0.00482)	(0.00509)
T × (30 Day Income)					0.0162				0.00771	0.00860
					(0.0125)				(0.0133)	(0.0135)
T × (Total Assets)						0.0144**			0.00694	0.00568
						(0.00700)			(0.00887)	(0.00884)
T × (Any Sex in Past 7 Days)							0.0160		0.00995	-0.00879
							(0.0353)		(0.0367)	(0.0611)
T × (Numeracy Score)								0.0223	0.0130	0.0166
								(0.0165)	(0.0191)	(0.0192)
Interactions w/other covariates‡	N	N	N	N	N	N	N	N	N	Y
Control-group mean	0.827	0.827	0.827	0.827	0.827	0.828	0.827	0.827	0.827	0.833
Control-group SD	0.264	0.264	0.264	0.264	0.263	0.264	0.264	0.264	0.264	0.259
Observations	1,282	1,282	1,282	1,267	1,197	1,281	1,282	1,282	1,182	1,160

Notes: All regresions include controls for stratification cells, interviewer fixed effects, and the alternate sexual behavior index from Appendix Table 1. Sample includes 1,292 sexually-active adults who were successfully interviewed at both baseline and endline; 120 of these have missing data for at least one of the covariates. Heteroskedasticity-robust standard errors, clustered by village, in parentheses: * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

‡ Other covariates include cognitive ability score, immediate word recall, any sex in the past 30 days, total sex acts in the past 7 days, an indicator for respondent changing their answers, and categorical indicators for ethnic group.