The Impact of Being Surveyed on the Adoption of Agricultural Technology

Mark Treurniet, Development Economics Group, Wageningen University & Research

Contact information: Mark Treurniet, PO Box 8130, 6700 EW Wageningen, Netherlands, telephone: +31 317 48 43 60, e-mail: mark.treurniet@wur.nl

* I thank Erwin Bulte and two anonymous reviewers for suggestions and comments, and Vivian Hoffmann, Sarah Kariuki and Janneke Pieters for the great collaboration in the wider research project. The data used in this article were collected in this project, which was supported by funding from the Netherlands Organisation for Scientific Research, through the Food & Business Global Challenges Programme. Data are provided through Dataverse at https://doi.org/10.7910/DVN/FQCOEK.

Include the DOI when citing or quoting: https://doi.org/10.1086/717251.

Abstract

This paper uses exogenous variation in the probability of being surveyed at baseline to estimate the impact of being surveyed on subsistence farmers' take-up of a new agricultural technology that improves food safety. I find large and statistically significant impacts of being surveyed, and also find that an experimental treatment effect disappears for surveyed farmers. My results have strong implications for our understanding of the process of technology adoption, for the external validity of adoption results measured in surveyed populations, and for research ethics.

1. Introduction

While the adoption of new technologies is an important driver of growth in output and quality of agricultural production, adoption has remained low in many developing countries (Foster and Rosenzweig 2010). Many studies therefore explore explanations for this low agricultural technology adoption (Foster and Rosenzweig 2010, Jack 2013). Often, such studies involve household surveys (e.g. Pamuk, Bulte, and Adekunle 2014, Aker and Ksoll 2016, Omotilewa et al. 2018, Omotilewa, Ricker-Gilbert, and Ainembabazi 2019). Yet, the effects of such surveys on later technology adoption and the estimation of parameters of interest are understudied.

Being surveyed may affect subsequent behavior in several, possibly related ways. First, survey questions may reveal new information to respondents. Second, surveys may focus scarce cognitive capacity and executive control to issues raised in the survey. Third, question-behavior effects arise if answering questions on predictions or intentions on specific behavior affects behavior (Rodrigues et al. 2015). Fourth, experimenter demand effects emerge if surveys change the respondent's perception of what the experimenter regards as "appropriate" behavior and this affects decisions (De Quidt, Haushofer, and Roth 2018). Fifth, being surveyed may make respondents aware that their behavior is observed as part of a study. Hawthorne effects arise if this awareness changes later behavior (McCarney et al. 2007).

The contribution of this paper is twofold: Substantively, studying the impact of being surveyed helps to better understand the process of agricultural technology adoption. Methodologically, studying how baseline surveys affect outcomes and experimental treatment effects helps to judge the external validity of results.

Substantively, while market inefficiencies have often been considered as constraining factor for agricultural technology adoption (Foster and Rosenzweig 2010, Jack 2013), much less attention has been given to mental processes and behavioral responses that might be triggered by surveys. While previous studies find that questions-behavior effects and experimenter demand effects are generally limited (Rodrigues et al. 2015, De Quidt, Haushofer, and Roth 2018), Schilbach, Schofield, and Mullainathan (2016) argue that the availability of scarce mental resources, or so-called "bandwidth", may be especially important for technology adoption processes. This paper therefore studies the impact of being surveyed on technology adoption among subsistence farmers.

Methodologically, this paper relates to a larger literature on panel conditioning. In a review paper, Warren and Halpern-Manners (2012) define panel conditioning as "bias introduced when participating in one wave of a longitudinal survey changes respondents' attitudes and behaviors and/or the quality of their subsequent reports of those attitudes and behaviors". Relative to most of the existing literature reviewed by Warren and Halpern-Manners (2012), this paper has three distinguishing features: First, many studies rely on self-reported outcome data, and thus cannot distinguish between behavioral effects and reporting effects. Since I use outcomes derived from administrative data, I am able to rule out the latter. Second, while researchers have rarely studied how surveys affect the estimation of experimental treatment effects, I am able to explore the effect on a cross-cut experimental treatment. Third, to the best of my knowledge, this paper is the first paper that studies the impact of being surveyed on the adoption of agricultural technology.

Methodologically, this paper is closest to Zwane et al. (2011), who report five different field experiments on the impact of being surveyed about health and/or household finances. In three health experiments, they find that being surveyed increases the use of water treatment products and take-up of medical insurance. In two microfinance experiments, they do not find an effect of being surveyed on borrowing behavior. The authors speculate that these results can be explained by what Schilbach, Schofield, and Mullainathan (2016) later defined as bandwidth, which was argued to limit the adoption of water treatment products and medical insurance more than the adoption of microfinance.

The objective of this paper is to explore the external validity of the results of Zwane et al. (2011), and study the impact of being surveyed on the take-up of a new agricultural technology that improves food safety. As part of a larger baseline survey, questions were asked on a specific food safety issue and experience with preventative measures, but no questions were asked on predictions or intentions to adopt a newly available technology. I use randomized variation in the probability of being surveyed at baseline to find large and statistically significant impacts of being surveyed on both the extensive and intensive margin of adoption, as recorded for both surveyed and non-surveyed farmers during sales that were organized by the research team. Since I use this administrative data, I am able to isolate effects on real economic behavior from effects on self-reporting. The evidence suggests that the increase in adoption is a direct result of being surveyed, but remains inconclusive about the contributions of bandwidth, experimenter demand effects and Hawthorne effects. Moreover, I find that the experimental treatment effect in a related experiment disappears for surveyed respondents.

The remainder of this paper is structured as follows: The next sections discuss the context in which the impact of being surveyed is studied, the survey experiment and the available data. Subsequently, I discuss the empirical strategy and results, first for the impact of being surveyed, and then for the interaction with related parameters. The last section discusses implications.

2. Study context

Aflatoxin is a fungal toxin that can grow in maize and groundnuts. High levels of exposure to aflatoxin may cause cancer, liver damage and death (Strosnider et al. 2006). Effective technologies to reduce aflatoxin contamination exist, but as with other agricultural technologies, their up-take has remained low.

This study focuses on Meru, Embu and Tharaka Nithi counties in the Eastern region of Kenya, which are known for their high levels of aflatoxin contamination. I study the impact of baseline surveys on the adoption of Aflasafe™, a biocontrol product that effectively decreases aflatoxin contamination in maize, and was recently introduced in Kenya. Aflasafe™ needs to be applied two to three weeks before flowering, and protects maize during its growth and storage (Bandyopadhyay et al. 2016).

The farmers in our study are well-aware of the problem of aflatoxin contamination and ways to prevent it, but the biocontrol product Aflasafe™ is not widely known. Among farmers surveyed at baseline, a large majority has heard about aflatoxin before (88.7%) and can tell what aflatoxin is (79.5%). After being told that "Aflatoxin is a poison that is produced by mold on maize and other crops", most are also able to mention at least one aflatoxin-induced health effect (81.4%) and at least one way to prevent aflatoxin from affecting maize (81.4%), such as drying maize well before storage (75.2%), which most actually have done (74.1%). However, when explicitly asked, only a small proportion reports having heard about Aflasafe™ (10.4%), and very few have used Aflasafe™ before (1.9%).

In a related paper, Hoffmann et al. (2020) study the impact of providing a modest market premium for aflatoxin-safe maize on the adoption of Aflasafe™ among subsistence farmers. After a census of existing producer groups, 152 groups from 124 villages were selected into the study. All groups were offered an opportunity to buy Aflasafe™ at a 50% subsidized price, and trained on its use. In addition, half of the 152 groups were offered a modest market premium for each bag of aflatoxin-safe maize sold through the project. This market linkage treatment was randomized across villages. Hoffmann et al. (2020) find that offering the market linkage does not increase adoption at the extensive margin (whether the individual purchased Aflasafe™ or not), but does increase adoption at the intensive margin (the quantity of Aflasafe™ purchased in kg).

Using a model, Hoffmann et al. (2020) argue that the modest market premium, which for many farmers was below the below the expected cost of adopting Aflasafe™, was likely to be insufficient to trigger adoption among farmers with a low valuation for aflatoxin-safe maize for home consumption, but could increase the quantity purchased by farmers with a high valuation for safe home consumption by decreasing the expected value lost from selling excess safe maize to the market in years with a good harvest. These farmers then prioritize safe maize for home consumption and sell only excess safe maize to the market.

3. Survey experiment

Prior to project trainings and Aflasafe™ sales, a baseline survey was conducted between September and October 2017. The survey was conducted by staff of our partner Innovations for Poverty Action (IPA). The baseline survey included questions on household demographics, household and livestock assets, land ownership and use, agricultural input use, maize harvest, post-harvest handling of maize, maize marketing, maize sale, expectations for the coming season, income sources, risk and insurance, group membership, as well as some questions on knowledge of aflatoxin and experience with aflatoxin prevention measures.

Out of at most 158 questions,[1] at most nine questions directly relate to aflatoxin:[2] The first four questions elicit active knowledge of aflatoxin. The following two questions elicit active knowledge and experience with aflatoxin prevention measures. The survey then elicits passive knowledge of Aflasafe™ by asking whether the respondent has ever heard of Aflasafe™. If and only if the respondent answers positively on this question (10.4%), the survey asks whether the respondent has ever used Aflasafe™, and whether the respondent has been given of promised to use this season. While the survey questions do focus attention to the problem of aflatoxin, and communicate that a technology called Aflasafe™ exists, they thus provide minimal further information about Aflasafe™.

For budgetary reasons, not all farmers were selected for baseline surveys. The sampling procedure created random variation in the probability of being surveyed. First, six primary members per group were randomly selected for interviews. These farmers were called 2-3 days before the day of the surveys. Subsequently, six members were randomly selected to replace the initially selected members if interviews could not be completed. The order of individuals within the replacement list was randomized as well. If less than six surveys could be completed among primary respondents on the day of the surveys, enumerators started calling replacement respondents from the top of the replacement list.

While the sampling procedure created random variation in the probability of being surveyed, it also prioritized some members over others, for which I will control in the analysis. During the group census, which had taken place between April and August 2017, a list was made of

[1] The survey includes several skip-criterions. The number of questions that a respondent received thus depends on his/her answers. 158 is the maximum number of questions a respondent can receive.

[2] The questions on knowledge of aflatoxin and experience with aflatoxin prevention measures are included in the Appendix.

all group members. Baseline survey respondents were first selected from members that had been present during the group census meeting. When this pool was exhausted, respondents were subsequently selected from members that had not been present during this meeting. For practical reasons, only the first twenty present members listed within each group were included in the randomization, and the remaining members were not considered. If the pool of those present at the meeting was exhausted before the required twelve members were selected, then only the first twenty non-present members listed were included in the randomization of selection of the remaining respondents, and the remaining non-present members were not considered. In the analysis for this paper, I include only those individuals that were included in the randomization,[3] and the econometric specification controls for the prioritization of present members over non-present members. For ease of expression, I define as "subgroup" all farmers within a group that were present, respectively not present, during the group census meeting, so that within a subgroup all farmers in my sample have equal probability of being selected into my primary or replacement sample.

Table 1 reports a breakdown of the sample. Among the 1786 individuals that were present in the group census meeting, 888 individuals (49.7%) were randomized into the primary list, 573 individuals (32.1%) were randomized into the replacement list and 325 individuals (18.2 %) were randomized in the control list, and almost half of these 1786 individuals were actually surveyed. From the 786 individuals that were not present in the census meeting, most were assigned to the replacement and control lists, and most of these 786 individuals were therefore not surveyed. In total, 876 individuals from the primary and replacement lists were

[3] For one group, the group census list was lost and re-taken later, making it impossible to retrieve which farmers were considered for surveys and which farmers were not. I therefore exclude this group from this analysis. Since I use individual level variation within groups to identify the impact of being surveyed, excluding this one group does not affect the internal validity of the estimated impact of being surveyed.

surveyed and 1696 individuals were not surveyed, and the total sample consists of 2572 individuals.

Figure 1 shows the proportion surveyed by survey selection status. Being selected in the primary sample increases the probability of being surveyed at baseline by 0.727 as compared to not being selected for surveying, which by design leads to not being surveyed. As expected, being selected as replacement still increases the probability of being surveyed at baseline as compared to not being selected for surveying, but the effect decreases with the rank ($p = 0.000$). As primary respondents were called 2-3 days before the day of the surveys and necessary replacement respondents were only called at the day of the surveys, primary respondents were more likely to be surveyed than replacement respondents conditional on being called ($p = 0.000$). The conditional probability of being surveyed does not significantly differ among replacement respondents with different rank ($p = 0.3651$).

4. Data

During the group census meeting, which took place between April and August 2017, all members of each group were listed, as well as whether they were present during this census meeting. A proxy for the farmer's gender could be derived from their names. As could be expected from the randomized setup, gender is not correlated with treatment assignment ($p = 0.446$, $p = 0.866$ and $p = 0.984$, respectively, using the main specification and robustness checks as discussed below with gender as dependent variable).

The baseline survey took place between September and October 2017, and was conducted by temporary IPA staff. The resulting data is available for surveyed respondents only, so it is impossible to use survey data to conduct further balance tests and empirically verify that individuals across the randomized primary, replacement and control lists have similar baseline characteristics.

The sale of Aflasafe™ took place between November and early December 2017, a few weeks after planting and just before Aflasafe™ should be applied, and was organized through a group meeting by staff of the Cereal Growers' Association (CGA). During sales meetings, newly hired staff from IPA[4] recorded the identities of buyers and quantities purchased, both for surveyed and non-surveyed respondents. I am therefore able to match the administrative sales data to the randomized sample selection and actual survey completion statuses. Farmers that wanted to buy less than one 4-kg package were requested to pair up with other farmers and share a package, but each farmer's share of the purchase was recorded. I derive two outcome variables from the sales data: Adoption measured at the extensive margin (whether the individual purchased Aflasafe™ or not) and adoption measured at the intensive margin (the quantity of Aflasafe™ purchased in kg).

Finally, baseline respondents were revisited in March and April 2018 for an endline survey. Since only baseline respondents were surveyed at endline, endline data variables cannot be correlated with being surveyed at baseline, and I therefore use the endline survey data only in some exploratory analysis.

5. Impact of being surveyed

5.1. Empirical strategy

To study the impact of being surveyed on the adoption of Aflasafe™, I estimated:

$$Adoption_{ig} = \beta Surveyed_{ig} + \gamma_g + \delta_g Present_{ig} + \varepsilon_{ig}, \qquad (1)$$

[4] While some baseline survey enumerators were hired again to record sales, they were re-assigned to groups, and only about 5% of the actual purchases of baseline respondents was recorded by their baseline survey enumerator.

where $Adoption_{ig}$ is the actual adoption of Aflasafe™ of individual i from group g, which I will analyze at both the extensive margin (whether the individual purchased Aflasafe™ or not) and the intensive margin (the quantity of Aflasafe™ purchased in kg), $Surveyed_{ig}$ is a dummy indicating whether the individual is surveyed at baseline, and ε_{ig} is an error term. As survey status was randomized within groups and respondents were first selected from present members and subsequently from non-present members, I included separate group fixed effects for present and non-present members $\gamma_g + \delta_g Present_{ig}$ to control for randomization strata. These fixed effects capture the variation between subgroups, so that the impact of being surveyed will be identified solely within subgroups. Note that the experimental treatment described in Hoffmann et al. (2020) was randomized at the village level (one level up from the group level), and is therefore captured in these group fixed effects.[5]

As there is partial compliance to the randomized survey status, actual survey completion is likely to be endogenous. I therefore instrument $Surveyed_{ig}$ by a dummy for selection into the primary sample and six dummies for selection as replacement, while expecting the probability of being surveyed to be highest for the primary sample and decreasing in the rank for replacements. More formally, I estimated:

$$Surveyed_{ig} = \lambda_0 Primary_{ig} + \sum_{r=1}^{6} \lambda_r \, Replacement_{igr} + \rho_g + \tau_g Present_{ig} + \omega_{ig}, \qquad (2)$$

where $Primary_{ig}$ is a dummy for selection into the primary sample, $Replacement_{igr}$ is a dummy for selection as r^{th} replacement and ω_{ig} is an error term. The fixed effects $\rho_g +$

[5] Additionally including the farmer's gender as covariate does not meaningfully affect my results.

$\tau_g Present_{ig}$ capture the variation between subgroups, so that the treatment effect will be identified solely within subgroups, where my instruments were randomized.

To test the strength of my instruments, I test H_0: $\lambda_0 = \lambda_1 = \cdots = \lambda_6 = 0$. The resulting F-statistic is 126.30 ($p = 0.000$). I thus reject the null hypothesis that the instruments are weak.

By using this instrumental variables design, I effectively estimate a Local Average Treatment Effect (LATE). Results were obtained using two-stage least squares (2SLS). Since groups were sampled from a larger population and the treatment effect might be heterogeneous, standard errors were clustered at the group level (Abadie et al. 2017).

5.2. Results

Table 2 reports results on the impact of being surveyed. I find large and statistically significant impacts of being surveyed. Adoption increases by 9.2%-point at the extensive margin and by 0.355kg at the intensive margin as compared to the control group. In line with the expectations expressed in the Introduction, being surveyed thus leads more farmers to adopt Aflasafe^{TM}.

To put these LATE estimates in perspective, I would ideally like to compare them with the relevant comparison, which is the mean of the adoption of the non-surveyed compliers to the randomized survey status. Since it is not possible to identify these compliers, I instead report the means of outcomes for the three lists: the control list, the replacement list and the primary list. Individuals on the control list were less likely to be present in the group census meeting, so that their adoption levels are probably lower than the relevant comparison, but individuals from the treatment list were more likely to be present in the group census meeting and are often surveyed, so that their adoption levels are higher than the relevant comparison. The relevant comparison is thus likely to fall between the mean of the control list and the mean of the primary list. Regardless of the comparison, however, the relative increase in adoption is substantial.

5.3. Robustness

The impact estimate presented above makes use of three lists: the primary list, the replacement list and the control list. Causal inference relies on comparability across lists within subgroups, which in expectation is created by randomization.

In a recent paper, however, de Chaisemartin and Behaghel (2020) show that waitlist estimators can be biased if the intended number of surveys can be reached before all individuals on the potential randomized waitlists reached an offer. In this case, one knows that the individual accepting the last offer by definition is a "taker", someone who would accept an offer to be surveyed if (s)he would receive it. Therefore, the expected probability of being a taker is higher among those that received an offer than among those that did not receive an over. If this probability of being a taker is correlated with adoption, the impact estimate is biased. As a first robustness check, I therefore estimate the doubly-reweighted ever-offer (DREO) estimator as proposed by de Chaisemartin and Behaghel (2020), which corrects for this bias by giving a lower weight to takers that received an offer. Results are presented in Table 3, Columns 1-2, and are similar to the main results.

Since individuals on the primary list always receive an offer, this issue would not exist if we compare individuals on the primary list with control individuals and exclude individuals from the replacement list from the sample. I therefore estimate the impact while excluding individuals on the replacement list from the sample. Results are presented in Table 3, Columns 3-4, and are slightly larger than the main results. A possible explanation for this difference can be that this effect is identified in a different subsample: The estimator relies on within-subgroup variation in the randomized survey status, which only exists in subgroups that are large enough to have individuals subscribed on all three lists. The sample sizes in parentheses show that only part of the groups have such within-subgroup variation in the randomized survey status.

5.4. Mechanisms

In this Subsection, I will discuss potential mechanisms that could have driven the impact of being surveyed on Aflatoxin™ adoption. While the impact on adoption could have been driven by the survey questions on aflatoxin knowledge and experience with aflatoxin prevention measures, surveys might also have affected adoption in more indirect ways. In this Subsection, I will first explore two potential indirect mechanisms. After I do not find empirical support for these indirect mechanisms, I will discuss which direct mechanisms could have driven the impact on adoption.

First, being surveyed could also have increased related production decisions that may affect the expected yield and the variance of the yield, which in turn could affect the adoption of Aflasafe™ (Hoffmann et al. 2020). I will therefore explore whether increased use of inorganic fertilizers or stress tolerant seeds can explain the impact of being surveyed.

Both the baseline and endline survey measured the adoption of inorganic fertilizer for maize production in the past short rains season. Among surveyed farmers that planted maize in both the previous short rains season (Oct 2016-Feb 2017) and the current short rains season (Oct 2017-Feb 2018), fertilizer adoption significantly increased by 10.7%-point ($p = 0.000$), from 57.5% to 68.2%. Since no endline survey data is available for individuals that were not surveyed at baseline, I am unable to check whether or not non-surveyed respondents were also more likely to use fertilizer. However, since fertilizer adoption is only weakly correlated with Aflasafe™ adoption ($r = 0.085$), even a 10.7%-point increase in the adoption of fertilizer would not seem to be a good explanation for the 9.2%-point increase in the adoption of Aflasafe™.

The surveys did not include any information or questions on stress-tolerant maize varieties or other ways to decrease the variance of the harvest, so it seems unlikely that being surveyed has affected Aflasafe™ adoption via increased adoption of variance decreasing technologies. I thus do not find empirical support for changes in related production decisions to explain the impact of being surveyed on the adoption of Aflasafe™.

Second, being surveyed could also have increased participating in project trainings, which in turn could affect the adoption of Aflasafe™. During the first project training, farmers received information about Aflasafe™, the rainfall index insurance that was sold together with Aflasafe™ [6] and, when relevant, the market premium for safe maize. The use of Aflasafe™ was demonstrated during a second training. As long as presence in the training is similar in treatment and comparison groups, and interaction effects between the survey and the training do not exist, then this does not affect the impact of being surveyed as reported in Table 2. However, respondents were being called for baseline surveys, and the first project trainings were often organized at the same location on the same day as the surveys. The information received through phone calls, and already being at the location of the trainings could both lead to increased presence in the training. If presence at the training subsequently increased the adoption of Aflasafe™, then the impact estimates in Table 2 could at least partly explained by increased participation in the project trainings. More specifically, the 2SLS estimate would capture:

[6] Aflasafe™ was offered in a bundle together with a rainfall index insurance product that covered the initial investment in case of bad rainfall conditions. Some groups also had the opportunity to buy Aflasafe™ without the insurance product, but most farmers bought Aflasafe™ with the insurance product.

$$E[\beta] = \frac{p_s\beta_s + (p_{t,1} - p_{t,0})\beta_t}{p_s} = \beta_s + \frac{p_{t,1} - p_{t,0}}{p_s}\beta_t, \qquad (3)$$

where p_s is the probability of being surveyed, β_s is the impact of being surveyed, $p_{t,1}$ is the probability of presence at the training given that the farmer is selected to participate in the survey, $p_{t,0}$ is the probability of presence at the training given that the farmer is not selected to participate in the survey and β_t is the impact of presence at the training; and I assume that $p_{t,1} \geq p_{t,0} > 0$ and $\beta_t \geq 0$. Project trainings explain part of the impact estimates in Table 2 if $p_{t,1} > p_{t,0}$ and $\beta_t > 0$.

If project trainings explain part of the impact, then I argue that it will be largest for primary respondents, and I use this to perform a robustness check: Individuals on the primary list were called 2-3 days before the day of the surveys, while necessary individuals on the replacement list were only called at the day of the surveys. The advance invitation for individuals on the primary list could help them to make themselves available, actually show up and participate in the baseline survey and trainings. Indeed, Figure 1 shows that, conditional on being called, individuals from the primary list were more likely to be surveyed. If $p_{t,1}$ and p_s are positively correlated, the relative effect $(p_{t,1} - p_{t,0})/p_s$ is larger for farmers on the primary list than for farmers on the replacement list. If presence at the training subsequently increased the adoption of Aflasafe™, then the 2SLS estimate should be higher for primary respondents. As a robustness check, I therefore estimated the 2SLS specification including the instrumented interaction between being surveyed and primary respondent status, so that I can compare 2SLS estimates across primary and replacement respondents.

Table 4 shows the results of this estimation. The interaction effect between being surveyed and being a primary respondent is not significantly different from zero and the point estimate is negative. I thus find no indication for the project trainings to explain the result in Table 2.

While within subgroups, the assignment to the primary and the replacement list is randomized, *actual* participation in the surveys is endogenous. An alternative explanation for the absence of an instrumented interaction effect in Table 4 could therefore be that primary respondents that were actually surveyed structurally differ from replacement respondents that were actually surveyed. For example, if older farmers are more flexible to participate, even if they are called at late notice, and if these older farmers have a low impact of being surveyed, then the LATE for replacement respondents would be lower than the LATE for primary respondents, which would negate the positive effect hypothesized above. I therefore explore whether any structural differences exist between primary and replacement respondents that were actually surveyed. Table 5 reports differences on a selection of baseline survey variables that was also used in Hoffmann et al. (2020). I do not find structural differences between primary and replacement respondents, except on the propensity for social learning dummy, which indicates whether the number of people the farmer discusses technologies with in the village is above the sample median. I thus find little empirical support for this alternative explanation.

Given that I do not find empirical support for indirect mechanisms to explain the impact of being surveyed on Aflasafe^{TM} adoption, I will now discuss which direct mechanisms may have contributed to this impact. These mechanisms were already introduced in the Introduction. First, while information provision cannot be ruled out as a mechanism, information provision was minimal. If this minimal information had driven Aflasafe^{TM} adoption, one might expect at least part of this effect to run via increased participation in project trainings, in which farmers received much more information. However, I did not find empirical support for this above. Second, bandwidth could have been an important mechanism, as at least seven questions focused attention on aflatoxin, its consequences and its prevention measures, mostly by stimulating active knowledge. Third, since no questions were asked on predictions or intentions to adopt aflatoxin prevention measures, question-behavior effects can be ruled out. Fourth, although questions were asked in a neutral way, the simple fact that these questions were asked by a researcher, could have been interpreted as a signal that the researcher considers aflatoxin prevention "appropriate". If this is the case, then experimenter demand effects may explain the impact of Aflasafe™ adoption. Finally, while project trainings and sales were led by another organization, being surveyed may have made respondents more aware that their behavior is observed as part of a study, so Hawthorne effects might also have played a role.

6. Interaction with experimental treatment

6.1. Empirical strategy

To explore how being surveyed can affect experimental parameter estimates, I study how the impact of the market linkage treatment studied in Hoffmann et al. (2020) varies across, respectively, the full sample of 2572 individuals used in this paper, the subsample of surveyed individuals and the subsample of non-surveyed individuals. Following Hoffmann et al. (2020), I use the following specification:

$$Adoption_{igv} = \alpha + \beta Treated_v + \varepsilon_{igv}, \qquad (4)$$

where $Treated_v$ is a dummy representing the market linkage treatment, which was randomized at the village level v, and ε_{igv} is an error term.[7]

As a second step, to deal with endogeneity in survey completion, I formally assess the interaction between being surveyed and the impact of the market linkage treatment using an instrumental variables approach. To be precise, I estimate:

$$Adoption_{igv} = \alpha Treated_v \cdot Surveyed_{igv} + \beta Surveyed_{igv} + \gamma_g + \delta_g Present_{ig} + \varepsilon_{igv}, \quad (5)$$

where variables are defined as before, while instrumenting $Surveyed_{igv}$ and $Treated_v \cdot Surveyed_{igv}$ by the primary respondent status, the six replacement respondent statuses, and their interactions with the market linkage treatment. Note that since market linkage treatment was randomized at the village level, $Treated_v$ is again absorbed by the group fixed effects γ_g in this specification.

This specification identifies how the impact of the market linkage treatment varies across the exogenous variation in being surveyed. Results were obtained using 2SLS and standard errors were clustered at the village level.

6.2. Results

Table 6 reports estimates of the impact of the market linkage in, respectively, the full sample of 2572 individuals used in this paper, the subsample of surveyed individuals and the subsample of non-surveyed individuals. Columns 1-3 report results for the extensive margin, and show no significant impact in any of the three samples. However, Columns 4-6 show that

[7] Hoffmann et al. (2020) also estimate regressions with a vector of baseline controls. Adding these baseline controls provides similar results.

the estimated impact on the intensive margin (as measured by the quantity of adoption) is much smaller and statistically non-significant among surveyed respondents, while it is statistically significant in both the full sample and the subsample of non-surveyed individuals. Conclusions on the impact of the market linkage treatment thus critically depend on which sample is used in the analysis. This pattern is in line with the result in Zwane et al. (2011) that an experimental treatment effect disappeared where being surveyed improved outcomes.

To formally assess the interaction between being surveyed and the impact of the market linkage treatment, Table 7 reports the results on the interaction between the market linkage treatment and being surveyed. The size of the interaction effect in Column 2 of Table 7 is similar to the difference in the treatment effects reported in Columns 5 and 6 of Table 6, which implies the differential impact across those surveyed and those not surveyed can be traced back to exogenous variation in the probability of being surveyed. The differences across Columns 5 and 6 of Table 6 is thus not the result of heterogeneity in impacts across two endogenously different groups. However, while the interaction estimate is large, the standard error around the interaction estimate is large too, so that the interaction effect is statistically insignificant. I thus cannot rule out that the difference in impact estimates across surveyed and non-surveyed individuals is the result of natural variation across subsamples.

6.3. Potential mechanisms

Being surveyed could potentially affect the impact of an experimental treatment in direct and indirect ways. First, being surveyed could directly affect the impact of the treatment, as modelled by β in equation (4). However, in this study, the baseline survey does not contain any information or primes about the market linkage, about price premiums for aflatoxin-safe maize or about market sales in general, so it does seems unlikely that a direct effect explains the difference observed in Table 6.

Second, being surveyed could indirectly affect the impact of the treatment by affecting the context in which the experimental treatment operates. Section 5.3 did not find support for related production decisions to play an important role. However, as becomes clear from the bottom row of Table 6, in the no market linkage condition, adoption was much higher among surveyed individuals. If the baseline survey increased the farmers' valuation for safe maize for home consumption, then the additional motivation from the market linkage might simply have been less important.

7. Discussion

This paper finds large and significant impacts of being surveyed on subsistence farmers' adoption of a new agricultural technology that improves food safety. While the evidence presented in this paper suggests that the increase in adoption is a direct effect of being surveyed, the evidence remains inconclusive about the contributions of bandwidth, experimenter demand effects and Hawthorne effects.

Following Zwane et al. (2011), my findings have crucial implications for the study of technology adoption. Substantively, as already suggested by Schilbach, Schofield, and Mullainathan (2016), bandwidth may be an important factor for the study of technology adoption. If insufficient bandwidth (cognitive capacity and executive control) is dedicated to a specific issue, people might ignore new technologies that address this issue. This paper, however, focused on one aspect of technology adoption: the purchase decision. It is not self-evident that farmers who acquire a technology will also apply the technology properly. For example, Kariuki et al. (2020) experimentally vary the level of support with the application of Aflasafe™ and find that providing additional support on top of a one-off training increases the probability that farmers apply Aflasafe™ at the correct stage from 80% to 100%. Further studies can shed more light on the role of bandwidth in technology purchase and application processes.

Methodologically, experiments that rely exclusively on samples that are surveyed before outcomes are measured are likely to provide adoption estimates that are higher than they would be in non-surveyed populations. Moreover, if survey effects and other treatment effects are not additively separable, estimates of treatment effect for surveyed samples may not be valid for external populations. Biases may especially arise in situations where available bandwidth is an important driver of technology adoption, and financial costs and benefits play a smaller role. If controlling for baseline covariates is still preferred in such settings, and baseline covariates are likely to be correlated within groups, one can consider to survey a subset of the members of each group and use group level means of baseline covariates in the analysis of adoption among the remaining group members.

Ethically, while the efficacy of new technologies may not yet have been tested outside controlled agronomic experiments, researchers directly affect the adoption of technologies. If this investment does not pay-off, welfare is affected negatively. Moreover, bandwidth is a scarce resource and using some bandwidth for one task may leave less for other tasks (Mani et al. 2013, Bulte et al. 2014). This might lead to worse over-all outcomes. Further research should shed more light on such undesirable side-effects.

If bandwidth would not have been an important mechanism, and the substantial impacts on technology adoption should be fully attributed to experimenter demand effects or Hawthorne effects, implications could be much more severe for the interpretation of many more studies on technology adoption, especially since baseline surveys might be just one of many sources of experimenter demand effects and Hawthorne effects. If the simple fact of being surveyed would already create substantial experimenter demand effects or Hawthorne effects on peoples' actual purchasing decisions two months later, one may start to question the external validity of a large empirical literature on technology adoption.

References

Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey Wooldridge. 2017. "When Should You Adjust Standard Errors for Clustering?" *National Bureau of Economic Research Working Paper Series* No. 24003. doi: 10.3386/w24003.

Aker, Jenny C., and Christopher Ksoll. 2016. "Can mobile phones improve agricultural outcomes? Evidence from a randomized experiment in Niger." *Food Policy* 60:44-51. doi: https://doi.org/10.1016/j.foodpol.2015.03.006.

Bandyopadhyay, R., A. Ortega-Beltran, A. Akande, C. Mutegi, J. Atehnkeng, L. Kaptoge, A.L. Senghor, B.N. Adhikari, and P.J. Cotty. 2016. "Biological control of aflatoxins in Africa: current status and potential challenges in the face of climate change." *World Mycotoxin Journal* 9 (5):771-789. doi: 10.3920/wmj2016.2130.

Bulte, Erwin, Gonne Beekman, Salvatore Di Falco, Joseph Hella, and Pan Lei. 2014. "Behavioral Responses and the Impact of New Agricultural Technologies: Evidence from a Double-blind Field Experiment in Tanzania." *American Journal of Agricultural Economics* 96 (3):813-830. doi: 10.1093/ajae/aau015.

de Chaisemartin, Clément, and Luc Behaghel. 2020. "Estimating the Effect of Treatments Allocated by Randomized Waiting Lists." *Econometrica* 88 (4):1453-1477. doi: 10.3982/ecta16032.

De Quidt, Jonathan, Johannes Haushofer, and Christopher Roth. 2018. "Measuring and Bounding Experimenter Demand." *American Economic Review* 108 (11):3266-3302. doi: doi: 10.1257/aer.20171330.

Foster, Andrew D., and Mark R. Rosenzweig. 2010. "Microeconomics of Technology Adoption." *Annual Review of Economics* 2 (1):395-424. doi: 10.1146/annurev.economics.102308.124433.

Hoffmann, Vivian, Sarah Kariuki, Janneke Pieters, and Mark Treurniet. 2020. "Safe food for me – and maybe for you: Upside risk, premium market access, and producer demand for a food safety technology." Working paper, Social Sciences Group, Wageningen University.

Jack, B. Kelsey. 2013. Market Inefficiencies and the Adoption of Agricultural Technologies in Developing Countries. PAL (MIT) and CEGA (UC Berkeley): Agricultural Technology Adoption Initiative.

Kariuki, Sarah, Asha Bakari, Charity Mutegi, Ranajit Bandyopadhyay, and Vivian Hoffmann. 2020. "Efficacy of a Food Safety Technology in Farmers' Fields Under Varied Levels of Farmer Training." Working Paper, Social Sciences Group, Wageningen University.

Mani, Anandi, Sendhil Mullainathan, Eldar Shafir, and Jiaying Zhao. 2013. "Poverty Impedes Cognitive Function." *Science* 341 (6149):976-980. doi: 10.1126/science.1238041.

McCarney, Rob, James Warner, Steve Iliffe, Robbert van Haselen, Mark Griffin, and Peter Fisher. 2007. "The Hawthorne Effect: a Randomised, Controlled Trial." *BMC Medical Research Methodology* 7 (1):30. doi: 10.1186/1471-2288-7-30.

Omotilewa, Oluwatoba J, Jacob Ricker-Gilbert, and John Herbert Ainembabazi. 2019. "Subsidies for Agricultural Technology Adoption: Evidence from a Randomized Experiment with Improved Grain Storage Bags in Uganda." *American Journal of Agricultural Economics* 101 (3):753-772. doi: 10.1093/ajae/aay108.

Omotilewa, Oluwatoba J., Jacob Ricker-Gilbert, John Herbert Ainembabazi, and Gerald E. Shively. 2018. "Does improved storage technology promote modern input use and food security? Evidence from a randomized trial in Uganda." *Journal of Development Economics* 135:176-198. doi: https://doi.org/10.1016/j.jdeveco.2018.07.006.

Pamuk, Haki, Erwin Bulte, and Adewale A. Adekunle. 2014. "Do decentralized innovation systems promote agricultural technology adoption? Experimental evidence from Africa." *Food Policy* 44:227-236. doi: https://doi.org/10.1016/j.foodpol.2013.09.015.

Rodrigues, Angela M., Nicola O'Brien, David P. French, Liz Glidewell, and Falko F. Sniehotta. 2015. "The Question–Behavior Effect: Genuine Effect or Spurious Phenomenon? A Systematic Review of Randomized Controlled Trials with Meta-Analyses." *Health Psychology* 34 (1):61-78. doi: 10.1037/hea0000104.

Schilbach, Frank, Heather Schofield, and Sendhil Mullainathan. 2016. "The Psychological Lives of the Poor." *American Economic Review* 106 (5):435-40. doi: 10.1257/aer.p20161101.

Strosnider, Heather, Eduardo Azziz-Baumgartner, Marianne Banziger, Ramesh V. Bhat, Robert Breiman, Marie-Noel Brune, Kevin DeCock, Abby Dilley, John Groopman, Kerstin Hell, Sara H. Henry, Daniel Jeffers, Curtis Jolly, Pauline Jolly, Gilbert N. Kibata, Lauren Lewis, Xiumei Liu, George Luber, Leslie McCoy, Patience Mensah, Marina Miraglia, Ambrose Misore, Henry Njapau, Choon-Nam Ong, Mary T.K. Onsongo, Samuel W. Page, Douglas Park, Manish Patel, Timothy Phillips, Maya Pineiro, Jenny Pronczuk, Helen Schurz Rogers, Carol Rubin, Myrna Sabino, Arthur Schaafsma, Gordon Shephard, Joerg Stroka, Christopher Wild, Jonathan T. Williams, and David Wilson. 2006. "Workgroup Report: Public Health Strategies for Reducing Aflatoxin Exposure in Developing Countries." *Environmental Health Perspectives* 114 (12):1898-1903. doi: 10.1289/ehp.9302.

Warren, John Robert, and Andrew Halpern-Manners. 2012. "Panel Conditioning in Longitudinal Social Science Surveys." *Sociological Methods & Research* 41 (4):491-534. doi: 10.1177/0049124112460374.

Zwane, Alix Peterson, Jonathan Zinman, Eric Van Dusen, William Pariente, Clair Null, Edward Miguel, Michael Kremer, Dean S. Karlan, Richard Hornbeck, Xavier Giné, Esther Duflo, Florencia Devoto, Bruno Crepon, and Abhijit Banerjee. 2011. "Being Surveyed Can Change Later Behavior and Related Parameter Estimates." *Proceedings of the National Academy of Sciences of the United States of America* 108 (5):1821-1826. doi: 10.1073/pnas.1000776108.

Tables

Table 1. Sample sizes

		Surveyed	Not surveyed	Total
Present in group census meeting	Primary list	650	238	888
	Replacement list	172	401	573
	Control list	0	325	325
	Total	822	964	1786
Not present in group census meeting	Primary list	5	8	13
	Replacement list	49	247	296
	Control list	0	477	477
	Total	54	732	786
Total	Primary list	655	246	901
	Replacement list	221	648	869
	Control list	0	802	802
	Total	876	1696	2572

Table 2. Impact of Being Surveyed

	Outcome variables	
	(1)	(2)
	Adoption	Amount (kg)
Surveyed	0.092***	0.355***
	(0.036)	(0.113)
Estimator	2SLS	2SLS
Sample	Full	Full
Groups	151 (151)	151 (151)
Subgroups	235 (198)	235 (198)
Observations	2572 (2349)	2572 (2349)
Mean of control list	0.120	0.315
Mean of replacement list	0.223	0.625
Mean of primary list	0.299	0.901

Standard errors clustered at group level in parentheses
* p<0.10, ** p<0.05, *** p<0.01
Sample sizes in parentheses exclude subgroups with zero variance in the randomized survey status

Table 3. Robustness

	Outcome variables			
	(1)	(2)	(3)	(4)
	Adoption	Amount (kg)	Adoption	Amount (kg)
Surveyed	0.097***	0.294***	0.125***	0.487***
	(0.025)	(0.088)	(0.041)	(0.127)
Estimator	DREO	DREO	2SLS	2SLS
Sample	Full	Full	Primary & Control	Primary & Control
Groups	151 (128)	151 (128)	151 (74)	151 (74)
Subgroups	235 (128)	235 (128)	231 (74)	231 (74)
Observations	2572 (1760)	2572 (1760)	1703 (774)	1703 (774)
Mean of control list	0.120	0.315	0.120	0.315
Mean of replacement list	0.223	0.625	0.223	0.625
Mean of primary list	0.299	0.901	0.299	0.901

Standard errors clustered at group level in parentheses

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

Sample sizes in parentheses exclude subgroups with zero variance in, respectively, the ever-offer indicator variable (DREO estimator) and the randomized survey status (2SLS estimtor)

Table 4. Mechanisms

	Outcome variables	
	(1)	(2)
	Adoption	Amount (kg)
Surveyed	0.149*	0.563**
	(0.080)	(0.251)
Surveyed*Primary	-0.048	-0.178
	(0.063)	(0.207)
Estimator	2SLS	2SLS
Sample	Full	Full
Groups	151 (151)	151 (151)
Subgroups	235 (198)	235 (198)
Observations	2572 (2349)	2572 (2349)
Mean of control list	0.120	0.315
Mean of replacement list	0.223	0.625
Mean of primary list	0.299	0.901

Standard errors clustered at group level in parentheses

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

Sample sizes in parentheses exclude subgroups with zero variance in the randomized survey status

Table 5. Baseline Characteristics Across Primary and Replacement Respondents

	Primary			Replacement			Diff	
	N	Mean	SD	N	Mean	SD	P[1]	P[2]
Age of the farmer (completed years)	655	50.1	13.8	221	50.3	14.3	0.815	0.089
Years of education completed by head	655	7.14	4.04	221	7.19	3.95	0.876	0.948
Relationship with the head	655	0.586	0.493	221	0.584	0.494	0.947	0.400
Asset index	655	5.63	2.34	221	5.88	2.29	0.154	0.427
Total land under maize main season previous year (acre)	655	1.43	1.20	221	1.56	1.32	0.184	0.548
Maize harvest main season previous year (kg)	655	433	776	221	473	714	0.479	0.403
Maize marketing: whether sold any maize last season	655	0.467	0.499	221	0.493	0.501	0.504	0.295
Total expenditures on agr. inputs & labour main season previous year (KES)	655	10814	11874	221	11253	11063	0.617	0.645
Propensity for social learning dummy	655	0.498	0.500	221	0.416	0.494	0.035	0.048
Aflatoxin knowledge index	655	-0.029	0.811	221	0.032	0.744	0.306	0.175
Knowledge and experience with insurance	655	1.33	0.82	221	1.30	0.83	0.679	0.375
Individual trust index	655	-0.010	0.538	221	0.040	0.522	0.216	0.873
Qualitative risk aversion	655	-0.534	1.869	221	-0.489	1.916	0.758	0.974

[1] P-value based on difference between primary and replacement and robust standard errors

[2] P-value based on difference between primary and replacement after controlling for randomization stata, and robust standard errors

Table 6. Impact of Market Linkage Treatment in Different Samples

	Outcome variables					
	(1)	(2)	(3)	(4)	(5)	(6)
	Adoption	Adoption	Adoption	Amount (kg)	Amount (kg)	Amount (kg)
Market linkage	0.023	-0.024	0.041	0.218**	0.081	0.269***
	(0.029)	(0.045)	(0.028)	(0.098)	(0.166)	(0.086)
Estimator	OLS	OLS	OLS	OLS	OLS	OLS
Sample	Full	Surveyed	Non-surveyed	Full	Surveyed	Non-surveyed
Villages	123	123	123	123	123	123
Groups	151	151	151	151	151	151
Observations	2572	876	1696	2572	876	1696
Mean of no market linkage	0.206	0.348	0.137	0.518	0.965	0.298

Standard errors clustered at village level in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Table 7. Interaction with Market Linkage Treatment

	Outcome variables	
	(1)	(2)
	Adoption	Amount (kg)
Market linkage*Surveyed	-0.061	-0.215
	(0.073)	(0.215)
Surveyed	0.122**	0.460***
	(0.054)	(0.138)
Estimator	2SLS	2SLS
Sample	Full	Full
Villages	123 (123)	123 (123)
Groups	151 (151)	151 (151)
Subgroups	235 (198)	235 (198)
Observations	2572 (2349)	2572 (2349)

Standard errors clustered at village level in parentheses
* p<0.10, ** p<0.05, *** p<0.01
Sample sizes in parentheses exclude subgroups with zero variance in the randomized survey status

Figures

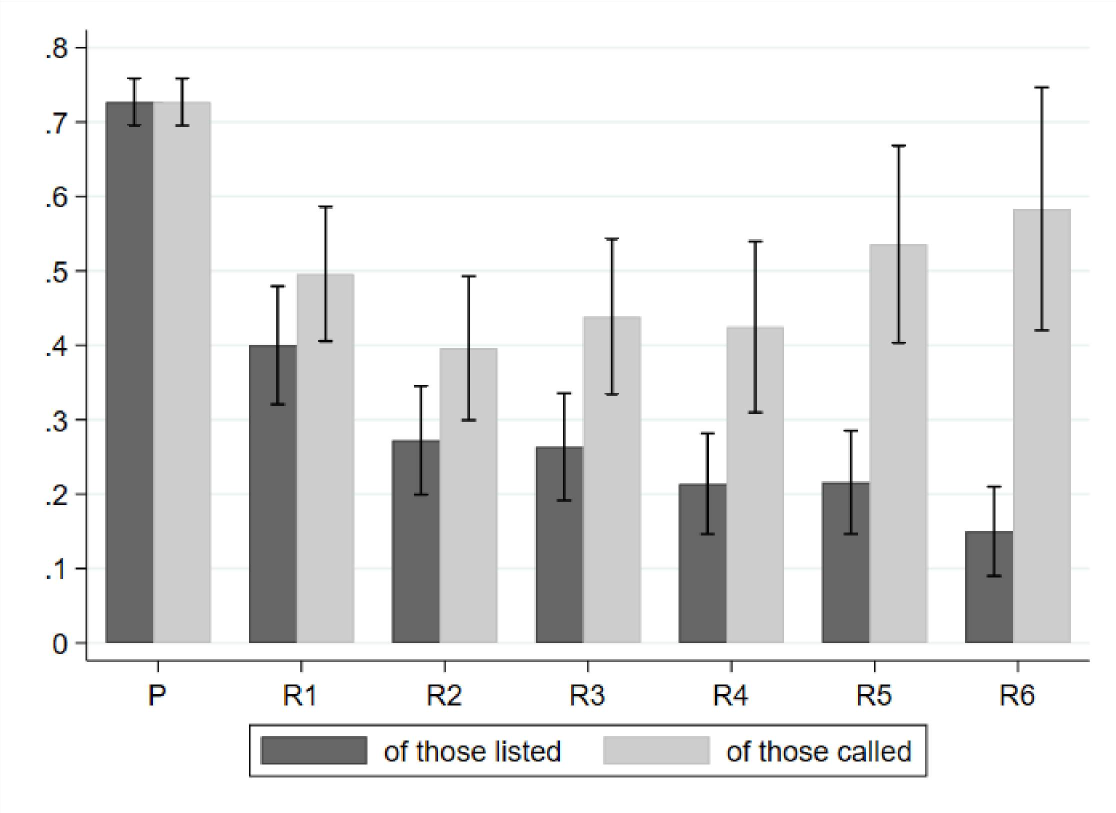

Figure 1. Proportion of individuals that are surveyed at baseline from the primary list (P) and the replacement list (R1-R6, where R1 refers to the first replacement and R6 refers to the sixth replacement)

Appendix: Survey questions on knowledge of aflatoxin and experience with aflatoxin prevention measures

	Question	Code
K.1	Do you know of any problem whereby eating maize can make you sick?	1=Yes 0=No
K.2	Have you ever heard of aflatoxin before today?	1=Yes 0=No
K.3	Can you please tell me what aflatoxin is? ***Do not read responses***	1=Mentions mold only 2=Mentions toxin / poison only 3=Mentions both 4=Maize disease 5=Does not know 96=Other specify
K.4	Read this before the respondent answers (regardless of whether they have heard of aflatoxin before): ***Aflatoxin is a poison that is produced by mold on maize and other crops*** Do you know any health effects that come from eating Aflatoxin? ***Record all responses mentioned*** ***Do not make suggestions***	1= Stomach Pain 2=Diarrhea 3=Lung problems 4=Jaundice 5=Liver failure 6=Cancer 7=Death 8=Stunting 9=Increases vulnerability to disease generally 99=Does not know any 96=Other specify
K.5	Do you know ways to prevent Aflatoxin from affecting your maize? ***Record all responses mentioned*** ***Do not make suggestions***	1=Drying maize well before storage 2=Drying maize off the bare ground 3=Storing the maize off the ground 4=Checking stores regularly for molds 5=Treat maize with fungicide before storage 6=Treatment of storage area before storage 7=Does not know 8=Use of bio control/ Aflasafe 96=Other specify
K.6	Did you take any measures to prevent Aflatoxin from affecting your maize ***Record all responses mentioned*** ***Do not make suggestions***	1=Drying maize well before storage 2=Drying maize off the bare ground 3=Storing the maize off the ground 4=Checking stores regularly for molds

		5=Treat maize with fungicide before storage 6=Treatment of storage area before storage 7=Does not know 8= Harvesting maize when it is fully mature 9=Use of bio control/ Aflasafe 96=Other specify
K.7	Have you ever heard of Aflasafe:	1=Yes 0=No –> Skip to K.10
K.8	Have you ever used Aflasafe	1=Yes 0=No
K.9	Have you been given or promised Aflasafe to use this season?	1=Yes 0=No