Be wary of those who ask: A randomized experiment on the size and determinants of enumerator effects*

Michele Di Maio†

Nathan Fiala‡

February 2018

Abstract

Enumerated surveys are popular because they allow for complicated questions to be asked in a fast and organized way, though there is little evidence of what effects enumerators can have on responses. We present evidence on this effect by randomly paring enumerators with respondents and find that there are likely minimal enumerator effects for less sensitive questions. However, we find large effects for very sensitive questions. Using data from a survey of the enumerators themselves, we find these effects can be captured by measurable enumerator characteristics. We argue that surveys with extremely sensitive questions should be designed with enumerator effects in mind.

JEL Classification Code: C81, C83, C93, O10

Keywords: Enumerator effects, survey experiment, randomized experiment

---

* We thank Areum Han and Mariajose Silva Vargas for excellent research assistance. We also thank the study participants for generously giving their time, as well as the field officers of Innovations for Poverty Action for their dedication to the quality of data collected. Data collection was funded by a Vanguard Charitable Trust and the World Bank. We are grateful for comments provided at various points of this study by Charles Towe, Harrison Diamond Pollock, Thoai Ngo and Jessie Pinchoff, and to participants of SITES 2017.

† Department of Business and Economic Studies, University of Naples Parthenope, Italy. E-mail: michele.dimaio@uniparthenope.it.

‡ Department of Agricultural and Resource Economics, University of Connecticut, USA. E-mail: nathan.fiala@uconn.edu.

1. Introduction

Data quality is a core concern of applied research. The ability of researchers to rigorously measure anything is determined by the reliability of the data collected. This is an increasingly important issue given the impressive growth in the number of researchers involved in data collection in the field and conducting empirical work using large micro-level datasets. Recently, the rising awareness of the significant implications of data quality for proper inference has led researchers to start looking at different issues that may affect quality, including the method of data collection, questionnaire design, and treatment of sensitive questions and hard-to-measure concepts (Beegle et al. 2012; Blair et al. 2012; McKenzie and Rosenzweig, 2012; Blattman et al., 2016; Serneels et al., 2016; Laajaj and Macours, 2017)[1].

However, best practices of collecting individual level data remain often (painfully and costly) to be learnt in the field or from colleagues. In particular, there has been relatively little systematic analytical work testing the presence and possible consequences for measurement and empirical analysis of the enumerator effect; i.e. that respondent's answers may be affected by the behaviour and characteristics of the enumerator.

The process of data collection through the filling of a questionnaire by a third party, or enumerator, is a social interaction between an interviewer and a respondent. It is thus possible that any element that may affect social interaction could also impact the quality of the data collected during the interview, including both active or passive influences on a respondent. For instance, the respondent may be influenced passively by his perception of

[1] On questionnaire design, researchers have investigated different aspects, including the wording of the questions (Beaman and Dillon, 2012), the specific place of questions within the survey questionnaire (Karlan and Zinman, 2012), and the length and level of detail of the questionnaire itself (Kalton and Schuman (1982) for an early review; de Mel, et al. 2009; Beegle et al., 2012; Serneels et al., 2016).

the interviewer – that is, by the interviewer's observable characteristics – and actively by the interviewer's behaviour – such as the interviewer's attitude and personality. Understanding whether this interaction can have implications for data quality should thus be of paramount concern to anyone who conducts surveys or uses survey data.

We conduct a large-scale experiment in Uganda to test for whether the characteristics of enumerators affects how people respond to survey questions. The questions range in sensitivity from not at all (age, sex, marital status) to potentially more sensitive (consumption of alcohol and tobacco, assets and level of education) to very sensitive (opposition party preferences). The experiment was conducted during the data collection of a project that covered 1064 villages and 6,895 respondents. We randomized, within language group teams, which village a randomly paired enumeration team would visit to conduct interviews with eight preselected people within that village.

We begin by testing for the predictive power of enumerator effects on respondent answers. We find that there is likely no or very little enumerator bias in reporting of most of the low and medium sensitivity questions the enumerators ask. However, we find very large potential enumerator bias in the reporting of support of opposition parties, a very sensitive question in Uganda. Moreover, we find that biases introduced by enumerators is coming from measurable, though not necessarily easily observable, personality traits.

We then test for the determinants of the enumerator bias on political support. We find that whether the enumerator identifies as coming from an urban versus rural area is positively associated with openness to vote for the opposition parties, though not for the ruling party. This result is consistent with the political situation in Uganda: support of the opposition parties comes from mostly urban areas, while rural communities vote overwhelmingly for the ruling party. Moreover, our results indicate that respondent's answers are influenced by enumerator's experience, indicating that capabilities of the enumerator in eliciting the correct information from the respondent is a crucial characteristic that may have a large impact on the survey results, though our design does not allow us to say whether this effect produces more or less accurate results. Finally, we find that respondents express lower support to both ruling and opposition parties to male enumerators, suggesting that female enumerators may be perceived as less concerning in expressing (sensitive) political opinions. All these results suggest that several individual enumerator characteristics may have important implications for respondent reporting for sensitive questions.

We end by discussing how these biases may affect data collection. For surveys with sensitive questions that are conducted with a specific identification strategy in mind, such as randomized and non-randomized impact evaluations, we argue that efforts should be made to ensure an enumerator interviews equal numbers of treated and non-treated individuals to balance out potential biases. This follows from the observation that some of the enumerator characteristics that we find to affect the respondent's answer (i.e. place of residence) there is no a clear a priori belief on its possible role or direction of the effect. This implies that balancing enumerators between treated and control is the only way to minimize enumerator bias. Surveys where an identification strategy may be identified in the future, such as census or living standards surveys, should include enumerator dummies in the data set for future researchers to conduct enumerator bias and balance tests. Future work could also explore whether there are ways to eliminate enumerators from certain questions. For instance, having respondents report their answers to a question using a tablet directly, rather than reporting it to an enumerator.

The literature studying the specific role of the enumerator during face-to-face interviews draws from different disciplines, including economics, political science, statistical methods, anthropology, and psychology (for a survey, see West and Blom, 2016). The enumerator effect can materialize due to enumerator characteristics directly affecting the answers of the respondent (Brunton-Smith et al., 2017) or because of the enumerator behaviour affecting non-response rates (Couper and Grove, 1992; West and Olson, 2010; Randall et al., 2013).

Previous research has shown that survey responses are associated with enumerator characteristics such as gender (Flores Macias and Lawson, 2008; Huddy et al., 1997), religion (Blaydes and Gillum, 2013; Benstead, 2014), ethnicity (Adida et al., 2016), experience and personality traits (Jäckle et al., 2013), differences in social status with the respondent (Kane and Macaulay, 1993), or even physical attractiveness (Jæger, 2016). There is also evidence that the enumerator effect varies with the type of question, being more salient for questions concerning gender-related issues, religion, ethnicity, corruption, and law and order (Schaeffer, 1980; Baird et al., 2008; Himelein 2016; Laajaj and Macour, 2017). Also, previous studies for the US have shown that the enumerator effect is likely to affect responses related to political opinions, with its influence depending on enumerator characteristics such as race, behaviour, and political views (see for instance, Davis and Silver, 2003).

One of the main challenge for this literature is to avoid confounding interviewer and respondent characteristics. The only way to exclude this possibility is to randomly assign interviewers to respondents. However, practical factors often prohibit the use of this design. Thus, much of the literature on interviewer effects consists of either telephone surveys with small numbers of interviewers or face-to-face surveys with no-random assignment of interviewers (West and Blom, 2016). Only a few studies have rigorously documented the causal impact of the enumerator on survey responses. These are a set studies on the race-of-interviewer effect (notably Williams, 1964; Reese et al., 1986; Cotter et al. 1982, Davis, 1997; Davis and Silver 2003) and on the gender-of-the interviewer effect (Catania et al. 1996; Huddy et al. 1997). Notably, all these studies have been conducted in the United States. Studies of the enumerator effect in developing countries are few (Flores-Macias and Chappell Lawson, 2008; Benstead, 2014, and Carlson, 2016) and all present a quasi-experimental design due to problems with randomization in the field. The only two exceptions are Himelin (2016) and Blayedes and Gyldum (2013). The former explore the presence of an enumerator bias in responses to a survey on corruption, women rights and community values in Timor Lest. The second looks at the effect of having an interviewer who wears a *hijab* on responses to questions related to religiosity and Islamic cultural norms in a survey experiment in Cairo (Egypt).

This paper presents four contributions to the literature on survey experiments and enumerator bias. First, we randomize enumerators to respondents, and do so across a large sample size, much larger than in previous studies. We randomize the pairing of 47 enumerators across four survey teams working in 1,064 communities and interviewing 6,895 individuals across the entire northern half of Uganda. This provides us with an unusually large sample for this type of experiment. Second, we collect detailed information on the enumerators using a survey that captures their demographics, work history, and a range of psychological measures. We show that the enumerator effects we observe can be captured from measurable indicators of enumerator characteristics. Moreover, our data allows us to explore the extent to which responses vary depending on whether the characteristics of the enumerator and the respondent match. Third, in our analysis we consider a large set of survey questions, ranging from not sensitive to very sensitive. We are thus able to test for the presence of the enumerator effect in the responses to several questions that are most common in surveys in economics and political science. While other studies have tested whether specific characteristics have effects on questions related to these same aspects, i.e. wearing a hijab on answers to questions related to gender issues or race on questions related to voting, our study shows that there may be enumerator characteristics that are important and not obvious ex ante. Moreover, different from previous studies, we look at types of questions and are thus able to compare for the same sample of respondents the magnitude of the bias across responses to questions that have different degrees of sensitivity.

Finally, while previous studies of the enumerator effect have focused on non-response rate, cooperation rates, or on differences in response by different individuals, we instead focus on the enumerator bias in actual responses. In particular, we study in detail responses to sensitive questions such as those related to political opinion, an issue of increasing importance in developing country studies.

To the best of our knowledge, this is the first paper to consider the causal effect of the enumerator bias on political opinion in a developing country. Our paper provides evidence of the existence of an enumerator bias, shows extent to which this may affect data quality for different outcomes, and provides recommendations and insights for future surveys. While previous studies have looked at how the quality of survey data depends on the characteristics of the survey questions, our analysis suggests that at least as much attention should be devoted to understanding the sometimes invisible, yet crucial, role of the enumerator during the interviews.

The remainder of this paper is structured as follows. In Section 2 we present the experimental design conducted on the enumeration team. In Section 3 we discuss the data, while the results are presented in Section 4. We conclude with a discussion of the implications of our results in Section 5.

2. Experimental design

The enumerator experiment described in this paper was done during the data collection of a separate research project, described in detail in Fiala and Premand (2018). That project was an experimental evaluation of a large-scale local accountability training program conducted in 2016 that included 1,064 villages and 8,403 respondents. Due to the dropping of some enumerators from our analysis, described below, our analysis here is on 6,895 interviews.

The research on the local accountability project included several economic and political outcomes, and so the questionnaire to respondents asked a range of sensitive questions, including age, sex, marital status, consumption of alcohol and tobacco, assets, level of education and opposition party preferences.

During the data collection, we randomized, within survey teams, which village a randomly paired enumeration team would visit to conduct interviews. We conduct the randomization on 47 enumerators and four teams covering the four-main language and geographic regions of northern Uganda.

We also conducted a survey of the enumerators that participated in the data collection. The survey collected information on enumerator demographics, including age, sex, whether their home is in an urban or rural area and education level. We also collected information on individual behavioural and psychological preferences. The survey was administered after the enumerators were selected to be part of the survey team, was voluntary, and was covered under the main project IRB.

Randomization was conducted such that pairs of enumerators were randomly chosen each morning to go to a specific community to conduct a data collection. Each community was visited the day before by a mobilizer, who confirmed that pre-selected individuals will be available for interviews with enumerators the next day. When a mobilizer finished their mobilization for the day, he or she then created a set of packages that contained tracking forms. Each package contained households that were close to each other. In the morning of the day of data collection, the field manager or team leader gave one package to each enumerator that was randomly determined using pre-developed randomization lists. Randomization was stratified by distance to ensure that if an enumerator went to a far village the day before, he/she would go to a closer one the current day, and vice versa. That is, the field manager or team leader, before randomizing the packages, split them into two groups: near and far. An enumerator who went to a far community or household the day before was given a randomly selected package from the near group. This process was done for each enumerator, making sure the distance between survey groups is well balanced for each person. This ensured no one enumerator felt that they were being given only very far communities and thus traveling significantly more than their colleagues.

As is common in data collections, there were times when someone from the enumeration team could not complete all of the surveys, and so a team leader had to conduct the survey. In other cases, enumerators did not complete their initial contract and were either fired or moved to the survey audit team. As the team leader was not randomly selected, and the selection of enumerators that left or were moved meant they completed a small number of surveys, we drop all surveys in which the interviewer conducted less than 70 interviews[2]. This reduces our sample from 8,403 interviews to 6,895. As a robustness check, we perform our analysis using other thresholds and find similar results.

The data collection required that all the people selected within a community to be interviewed needed to be completed in one day. Given the length of the survey and that there were eight people per community to be interviewed, this meant that two enumerators were needed per community. An enumerator was thus paired with another enumerator when given a community to visit. We were not able to randomize which specific individual within a community an enumerator interviewed. This could present bias in our analysis if enumerators could systematically choose individuals to interview of a specific type. However, in the analysis to follow, we show that this is unlikely to be the case. First, Chi-squared tests indicate that respondent characteristics (age, gender, marital status, and education level) are not correlated with enumerator characteristics (gender, place of residence, experience). Moreover, for relatively easily observable characteristics, such as

[2] Seventy interviews is the 10th percentile of the distribution of the number of interviews by enumerator.

age, gender, and income level, we do not find systematic bias in responses, confirming that there was no selection on these variables (see results in Table 2 below).

3. Data

The data collected as part of this study is presented in Table 1 and includes data on the respondents and the enumerators. We report here a limited number of variables for the enumerators, but include the full list of questions asked in the Appendix 2.

The enumeration team was 35% male, with 92% reporting they had attended at least some university. There was very low variation in age, with the average enumerator being 28 years old. 61% of the enumerators self-identified as being from an urban area.

For respondents, they are on average 44 years old, with 49% being men. Only 19% were single, with 5.8 years of education on average. Households had on average 8.2 members, with 2.3 heads of cattle. 41% of the sample reported having consumed alcohol or tobacco in the last week. They report total spending on alcohol and tobacco, conditional on consuming any, of 3,800 USH (approximately $1.20).

The respondent measurements are also presented in Table 1. We describe them here in order of our expectations of the sensitivity of the question. Individuals in Uganda are generally happy to report their age, gender, marital status, years of education and household size. People are less likely, though depending on the conditions still open to, discussing their wealth (measured here through assets), their number of animals, and how much alcohol they consume. These questions are often thought of as a major reason for conducting enumerated interviews: a good interviewer can make someone comfortable about reporting such things (Blair, Czaja and Blair, 2014; Fink, 2006).

We end by discussing the four questions we consider to be the most sensitive. It is common for people to state in public their support for the ruling party, the National Resistance Movement (NRM), as this is highly encouraged by the government. However, people are generally much more hesitant to report support for opposition parties, which include the Democratic Party (DP), Forum for Democratic Change (FDC) and Uganda People's Congress (UPC). This is in part due to perceived and actual government action against supporters of these parties, including general harassment, arrests and, much less common but reported in some international press and believed by many opposition supporters, torture[3]. For each of these four parties, we asked respondents the following question:

> *People have different feelings about different political parties. I am going to read you a list of different political parties in Uganda. For each one, I am interested in your openness to vote for a strong candidate from this party if there were to be one.*
>
> *Again, everything you tell me is confidential and cannot be shared with anyone outside our team. Also, please feel free to tell me if there is a question that you do not want to answer and we can move on to the next question.*
>
> *You will have to answer with a number between 1 to 4. 1 means not open at all to voting for a strong MP candidate from a party, and 4 means very open.*
>
> *1) Not open at all*
> *2) Somehow not open*
> *3) Somehow open*
> *4) Very open*

The enumerators were instructed to read the entire question and ensure that no one else was listening to the conversation. Note that in the context of Uganda it is very difficult

[3] Uganda can best be characterized as a semi-authoritarian regime. There is significantly greater political competition at the local than the national level, though the ruling party, which initially came to power in a coup in 1987, holds 2/3 of parliament seats and the majority of local elected positions. Tripp (2010) provides a detailed discussion of Ugandan democracy.

within communities to observe political preferences without asking. 91% of the people interviewed responded to this question. This is a very high rate compared to similar questions in the literature. Non-responses to political questions are very common and in fact are one of the aspects most studied in the literature (see for instance, Blair et al. 2012). Average support for opposition parties on the above scale was approximately 1.9. This was half of the average value given to the ruling party.[4] Voting preferences are highly correlated within communities: stated preference for opposition parties has an intra-class correlation between 0.30 and 0.33.

4. Results

We next present the results on the effect of enumerator characteristics on respondent outcomes. We first look at how much variation in responses is accounted for from an enumerator fixed effects model and by enumerator characteristics. We then look closely at the political questions to explore what individual enumerator characteristics may matter for the bias we observe.

4.1 Measuring the enumerator bias

Our main method to identify potential enumerator bias and its magnitude is to look at how much enumerators themselves contribute to any variation in responses that we observe. We thus begin by testing for the predictive power of enumerator effects on respondent answers

[4] Note that from the answer to this question it is not possible to know which party the respondent has voted for. This helps to explain why the response rate is much higher than in usual survey on political opinions. Yet, this comes at the cost of asking a less precise question. However, it should be noted that this formulation of the question - because the respondent is invited to express a separate and not comparative judgment on each political party - reduces the possibility of an enumerator effect because the respondent's political view is not uniquely identified. In this sense, we should interpret our estimates as a lower bound of the true effect.

by examining the R^2 in an enumerator fixed effects regression, similar to Himelein (2016) and Laajaj and Macour (2017). Results are reported in Table 2. In column 1, we report the results from regressing each outcome of interest (such as respondent's gender, age, household assets, political support, etc.) on a constant and dummies for each enumerator, with no other controls. A high R^2 is interpreted as the enumerator effect picking up a large amount of the variation in responses, while a low R^2 indicates that there is no or very low enumerator bias. We report adjusted R^2 to avoid artificial inflation of the R^2 value due to adding in additional variables. We report the average adjusted R^2 across all the teams.[5]

We find very low values for most of the questions, including whether the respondent is single (0.004), respondent gender (0.009), age (0.011), education level (0.012), household size (0.027), asset index (0.025), the number of cattle a household owns (0.027), and even whether the individual consumes and how much they spend on alcohol (0.009 to 0.023). These low values suggest two conclusions. First, individual enumerators do not systematically impact the way respondents report these answers. Second, it confirms that our enumerator randomization strategy worked well. In line with the results of the tests discussed in Section 2, the low value associated to respondent characteristics indicate that enumerators are not systematically choosing respondents within villages based on age, gender, education level, etc.

The adjusted R^2's for political party support, however, are substantially different than those for the other questions. We find a low value for NRM support, but the adjusted R^2 for opposition parties is high (between 0.234 and 0.292) and varies by enumeration team

[5] In appendix Tables A1-A4 we report the results by team for each column of Table 2.

from a low of 0.137 to a high of 0.434 (see Table A1). These values are very high and are not suggestive of random noise or unsystematic reporting of support.

We next explore whether the measurable characteristics of enumerators is causing this bias. In Table 2 column 2, we report the adjusted R^2 for a regression on the respondent response where, instead of including the enumerator fixed effect, we use a set of observable characteristics of the enumerators. These characteristics are age, gender, education level, and whether the enumerator's locality of residence is urban or rural area. The adjusted R^2 for all questions are relatively low, even for the political questions (from 0.001 to 0.173) and – for each question – they are lower with respect to those in Table 2 column 1. While observable enumerator characteristics shows some bias, the results suggest that including only few enumerator characteristics would mask the true bias.

In column 3, we then add to the same regression controls for the working ability of the enumerator. These are: (1) months of experience as enumerator; (2) how important an enumerator considers that people think he is good at his work; and (3) how much she/he is motivated by the money she /he can earn working as enumerator. The adjusted R^2 increases in size considerably compared with the values observed in column 2 (from 0.002 to 0.275). This indicates that including more enumerator characteristics can get closer to the true size of the enumerator bias.

Finally, we add to the basic enumerator characteristics a set of personality traits and present the results in column 4. Comparing the adjusted R^2 from these regressions with those in column 1, 2, and 3, we find that for all outcomes the value is much larger and - for the most part - identical to those found in column 1 using only fixed effect dummies. We conclude that it is possible to capture the complete bias from enumerators through individual, measurable enumerator characteristics.

4.2. Determinants of enumerator bias

The next step in our analysis is to look at what specific enumerator characteristics may be affecting individual responses. To this end, we focus on the questions in which the enumerator effect seems to matter most, namely voting preferences. Specifically, we estimate the following OLS regression model:

$$Y_{ic} = \alpha + \beta X_j + \gamma P_{jc} + \delta R_i + c + \varepsilon_i \tag{1}$$

where Y_{ic} is the outcome of interest for individual i, namely the openness to vote for each party, living in community c. X_j is a matrix of basic enumerator j characteristics (gender, age, education level and whether the place of residence is urban). P_j is a matrix of enumerator j additional characteristics related to ability and psychological characteristics. Finally, c is the set of community fixed effects and ε_{ic} is the error term. All standard errors are clustered at the enumerator level. As a robustness check, we also include a set of characteristics for respondent i living in community c captured in vector R_{ic} (age, gender, education level, and marital status).

In Table 3, we present the regressions results on stated support for different political parties for our baseline specification, i.e. only including basic enumerator characteristics. Some interesting results emerge. The coefficient for whether the enumerator is a man is always significant and negative, even though the magnitude of the coefficient is significantly smaller for the ruling party. The urban indicator is remarkably significant for all opposition parties. Respondents report significantly higher support for opposition candidates to urban enumerators. This effect is especially large when controlling for personality measures. There is no effect from urban status on stated support for the ruling party. In columns 5 to 8, we report the results when we include as control a set of ability and psychological questions that were asked of the enumerators. Including these additional variables increases the explanatory role of the enumerator being from an urban are and decreases that of the enumerator being male. While there is no clear pattern in the way enumerator psychological characteristics affect responses in the survey, we note that they are generally of a different sign between the three opposition parties and the ruling one.

4.3 Robustness checks

Our results are robust to various checks. First, all of our results are robust to dropping of enumerators that conducted less than 24 interviews (the bottom 1% of the enumerator's number of completed questionnaires), rather than less than 70 (as in our main analysis). Results - reported in Appendix Table A5 - do not change with respect to those obtained using our main sample. Second, our results for the effect of enumerator characteristics are robust to the inclusion of respondent characteristics such as age, gender, education level, and marital status (see Table A6). Third, we have re-done all the analysis for the political questions using as outcome a dummy variable rather than the continuous one. Results (reported in Table A7 available upon request) are unchanged.

Finally, we run a placebo test to show that enumerator characteristics do not affect all respondent answers, possibly capturing some unobserved common individual characteristic. To this end, we re-run model (1) using as alternative outcomes variables with different levels of sensitivity. These are: (1) the age of the respondent; (2) the household size; (3) the number of cattle owned by the household; and (4) if the respondent consumes alcohol or tobacco. Results reported in Table 4 show that enumerator characteristics do not explain responses regarding age, household size, and number of cattle owned by the respondent. The first two are clearly non-sensitive questions while the third is an easily verifiable question since it the main outcome of the program. On the contrary, enumerator characteristics such as age and gender do explain variation in the response to the use of alcohol or tobacco. In particular, respondents are more likely to report consumption to enumerators who are male and older. These results confirm our main finding indicating that enumerator characteristics do matter more for sensitive questions. At the same time, these results indicate that a priori it is not obvious which characteristic matter for each question.

5. Discussion

The results we present here suggest that individual enumerator characteristics may have important implications for respondents answering sensitive questions.

The first characteristic of interest is the enumerator being from an urban area rather than from a rural one. An extensive literature examines how social desirability concerns can influence the way survey or experimental participants answer questions. Social conformity and social desirability bias generally refer to the tendency of respondent to provide responses that she believes will be viewed favourably by others, anticipating the views of the enumerator and thus answer ways to please him or her. The respondent may express opinions that conform to societal norms or the interviewer's perceived expectations, implying that respondents' answers may be biased an effort to please the interviewer. These tendencies may be exacerbated on sensitive issues where fear and the desire to avoid embarrassment and criticism are stronger (Blaydes and Gillum, 2013). If this behaviour is influenced by the enumerator characteristics, this generates an interviewer effect. Social deference theory also suggests that a desire to minimize the social distance between two strangers may lead to responses that complement the interviewer's perceived social group. Differential social status and power is also believed to impact interview response bias (Lenski and Leggett, 1960; Davis 1997). In our case, the positive effect of the enumerator being from an urban center on the respondent being open to support an opposition party candidate may capture the fact that most opposition to the ruling party comes from urban centres[6]. While all the respondents in this sample are from rural areas, it is possible that they may become more open to expressing their support to opposition parties when in the presence of an enumerator from an urban area. Conversely, this may reflect a desirability bias in respondents and may not reflect their true preferences.

The second enumerator characteristic that turns out to be very significant in explaining the respondent answers is the number of months previously worked as enumerator, which we can be interpret as a proxy for the enumerator experience and/or ability. The strong effect of enumerator's experience is in line with evidence emerging from case studies of survey collection in sub-Saharan Africa presented in Randal et al. (2013). This indicates that the capabilities of the enumerator in eliciting the correct information from the respondent is a crucial (unobservable) characteristic that may have a

[6] See http://www.theeastafrican.co.ke/OpEd/comment/Museveni-NRM-party-still-has-huge-support-in-rural-Uganda-/-/434750/3036604/-/syo070/-/index.html for a discussion of the results from the 2016 election that occurred 3 months before this data collection and the role of rural voters in the NRM win.

large impact on the survey results. This is also in line with the results reported in Jäckle et al. (2013) who find evidence of a positive effect of experience on co-operation rates.[7] We must be cautious when interpreting this result though as our design does not allow us to identify whether experience produces less biased answers, or more biased.

Finally, we find a strong and robust effect of the gender of the enumerator on stated voting preferences of the respondent. This result is in line with several studies that have explored the role of gender in the context of survey data collection (see West and Blom, 2016). Our results show that respondents are less likely to report support for any party if the enumerator is male. Interestingly, for the ruling party this effect is smaller (meaning that respondent are more likely to report support for it with respect to other parties if the enumerator). In other words, if the enumerator is male (rather than female) the respondent is reporting a worse opinion for each party. This is in line with Axinn (1989), who suggests that female enumerators may be perceived as less frightening. Interestingly, this effect is smaller for the ruling party. Looking at respondent-enumerator gender interaction (see Table A8), we find that female respondents are significantly less likely to support the ruling party. The results suggest that responses may be biased not only by enumerator characteristics but also on how they interact with enumerator characteristics. These results add to the mixed results from the literature regarding interaction effects (see for instance Catania et al., 1996; Huddy et al. 1997; Himelein, 2015). While we do not explore in detail this additional possible source of enumerator bias, we suggest that this could be another interesting area for future research.

[7] They do not explore whether experience has a positive effect due to learning or through rather than selective dropout of less successful interviewers. To answer address this question, longitudinal data over several years would be needed.

6. Concluding remarks

Our results have shown that enumerator bias can be an important issue for the analysis of micro-level data. This type of measurement error can be of especially important concern if it is correlated with treatment status. In this case, the enumerator bias could affect both the size and the sign of the program impact. This bias could be even more problematic if combined with the possibility of a self-report bias.[8]

Understanding how enumerator characteristics may relate to respondent characteristics may influence data quality has important implications for the recruitment, selection, training and evaluation of interviewers and for the survey design. We now end by discussing what these results may mean for conducting surveys on potentially sensitive topics.

There are different possible solutions to the issue of enumerator bias that could improve the quality of enumerator data collection. Baird and Özler (2012) suggest using alternative sources of data (i.e. administrative data), though this is not always a viable option for some questions of interest. Recent research on list experiments (see for instance Blair and Kosuke, 2012) suggest they may produce more honest answers. However, these listing methods do not allow for individual level analysis. Future work on how to incorporate plausibly private reporting of information, such as handing an electronic data collection device to a respondent to complete a question, could be fruitful.

[8] Self-reported data can bias program impact in different directions. For example, comparing program impacts using self-reports vs. monitored data, Barrera-Osorio et al. (2011) report that a significant positive bias in self-reported school enrollment for all individuals compress the difference between the treatment and control groups causing a downward bias in observed program impact. Instead, Baird and Özler (2012) find evidence that differential misreporting does bias downward program impact in the case of an education program in Malawi because girls in a control group are more likely to over-report school enrolment compared to those receiving a conditional cash transfer.

In cases where such methods are not plausible, it is likely impossible to eliminate enumeration bias. Instead, there are ways to minimize such bias. For the case of randomized control trials, which this enumerator experiment was conducted in conjunction with, the simplest method is to ensure that enumerators equally interview both treatment and control individuals. Researchers can then report how much enumerator fixed effects or characteristics affect variation in treatment status. For example, as the enumerators were randomly assigned, the study used for this experiment has an R^2 on treatment status for respondents for observable enumerator characteristics of 0.002, observable plus personality characteristics of 0.020, and fixed effects of only 0.024. This suggests that there will be minimal if any bias from enumerators for sensitive questions for the identifying method used in this study (randomized assignment). More generally, researchers should provide evidence of enumerator balance around whatever identification strategy is used, especially when very sensitive questions are being asked.

Surveys that collect sensitive information and for which there is not a clear identification strategy but which may be used by other researchers in quasi or natural experiments, such as is done in censuses across the developed and developing world, Afrobarometer and the World Bank Living Standard Measurement Survey (LSMS), need to be interpreted cautiously. Teams that collect this data should provide, at a minimum, a test for enumerator fixed effects and a clear description of how enumerators and teams were assigned. Researchers can then provide balance tests, controls and bounding tests to account for potential enumeration biases. Simply including enumerator fixed effects in analysis is not likely to completely solve this issue unless assignment of enumerators is clearly documented and appropriately incorporated into the analysis.

To conclude, we note that, different from other improvements in measurement, our proposal does not involve significant additional survey costs and does not generate trade-offs between accuracy or bias and cost. On the contrary, for most studies, limiting enumerator bias can be done quickly, relatively easily, and inexpensively, and so represents a low-cost improvement in the quality of collected data.

References

Adida, C.L., Ferree, K.E., Posner, D.N. and Robinson, A.L. (2016). Who's asking? Interviewer coethnicity effects in African survey data. Comparative Political Studies, 1–31

Axinn, W.G. (1989). Interviewer and data quality in less developed setting. Journal of Official Statistics, 5, 265-280

Baird, S., Hamory, J. and Miguel, E. (2008). Tracking, attrition and data quality in the Kenyan Life Panel Survey Round 1 (KLPS-1). Center for International and Development Economics Research, Working Paper C08-151

Baird, S. and Özler, B. (2012). Examining the reliability of self-reported data on school participation. Journal of Development Economics, Volume 98, Issue 1, Pages 89-93

Barrera-Osorio, F., Bertrand, M., Linden, L. and Perez-Calle, F. (2011). Improving the Design of Conditional Transfer Programs: Evidence from a Randomized Education Experiment in Colombia. American Economic Journal: Applied Economics 3, no. 2: 167–95

Beaman, L. and Dillon, A. (2012). Do household definitions matter in survey design? Results from a randomized survey experiment in Mali. Journal of Development Economics 98, 124-135

Beegle, K., De Weerdt, J., Friedman, J. and Gibson, J. 2012. Methods of household consumption measurement through surveys: Experimental results from Tanzania. Journal of Development Economics, vol. 98(1), pages 3-18.

Benstead, J.L. (2014). Does interviewer religious dress affect survey responses? Evidence from Morocco. Politics and Religion, Volume 7, Issue 4, pp. 734-760

Blair, G. and Kosuke, I. (2012). Statistical analysis of list experiments. Political Analysis, 20(1), 40-77.

Blair, G., Fair, C., Malhotra, N., Shapiro, J. N. (2012). Poverty and support for militant politics: Evidence from Pakistan. American Journal of Political Science, Vol 57, 30-48

Blair, J., Czaja, R. and Blair, E. (2014). Designing surveys: A guide to decisions and procedures. Sage

Blaydes, L. and Gillum, R. M. (2013). Interviewer effects in the Islamic world: Evidence from a randomized survey experiment in Egypt. Politics and Religion, 6 (3), 459-485

Blattman, C., Gonwa, T., Jamison, J., Rodrigues, K., Sheridan, M. (2016). Measuring the measurement error: A method to qualitatively validate survey data. Journal of Development Economics, 120, 99–112

Brunton-Smith, I., Sturgis, P., and Leckie, G. (2017). Detecting and understanding interviewer effects on survey data by using cross-classified mixed effects location-scale model. Journal of the Royal Statistical Society A. Volume 180, Issue 2, 551–568

Couper, M.P. and Grove, R.M. (1992). The role of the interviewer in survey participation. Survey Methodology 18(2): 263-278

Davis, D. W., and Silver, B,D. (2003). Stereotype threat and race of interviewer effects in a survey on political knowledge. American Journal of Political Science. 47(1): 33-45.

Davis, D. W. (1997). The direction of race of interviewer effects among African Americans: Donning the Black Mask. American Journal of Political Science 41:309 – 322

de Mel, S., David McKenzie, M., Christopher Woodruff, C. (2009). Measuring microenterprise profits: Must we ask how the sausage is made? Journal of Development Economics 88, 19–31

Fiala, N., and Premand, P. (2017). Impact of social accountability on the quality of community projects in Uganda: impact evaluation results from NUSAF2 transparency, accountability and anti-corruption sub-component. Report ACS21095. Washington, D.C.: World Bank Group

Fink, A. (2006). To conduct surveys: A step-by-step guide. Sage.

Fisher, M., Reimer, J. J., and Carr, E. R. (2010). Who should be interviewed in surveys of household income? World Development, 38(7), 966–973.

Flores Macias, F., and Lawson,C. (2008). Effects of interviewer gender on survey responses: Findings from a household survey in Mexico. International Journal of Public Opinion Research, vol. 20 (1), 101-109.

Floyd, F. (2014). Survey research methods. Sage.

Himelein, K. (2016). Interviewer effects in subjective survey questions: Evidence form Timor-Leste. International Journal of Public Opinion Research, vol. 28 (4): 511-533

Huddy, L., Billig, J., Bracciodieta, J., Hoeffler, L., Moynihan, P.J., Pugliani, P. (1997). The effect of interviewer gender on the survey response. Political Behavior 19: 197-220

Jæger, M.M. (2016). Hello beautiful? The effect of the interviewer physical attractiveness on interview success. Sociological Methods & Research, forthcoming

Jäckle, A., Lynn, P., Sinibaldi, J., Tipping, S. (2013). The effect of interviewer experience, attitudes, personality and skills on respondent co-operation with face-to-face surveys. Survey Research Methods, Vol.7, No.1, pp. 1-15.

Kalton, G. and Schuman, H. (1982). The effect of the question on survey responses: A Review. Journal of the Royal Statistical Society. Series A (General), Vol. 145 (1), 42-73

Kane, Emily W., and Macaulay, L. J. (1993). Interviewer gender and gender attitudes. Public Opinion Quarterly 57:1–28

Karlan, D.S. and Zinman, J. (2012). List randomization for sensitive behavior: An application of measuring use of loan proceeds. Journal of Development Economics 98, 71-75

Laajaj, R. and Macour. K. (2017). Measuring Skills in Developing Countries. World Bank Policy Research Working Paper 8000.

Lenski, G. E., and Leggett, J.C. (1960). Caste, class, and deference in the research interview. American Sociological Review 65:463–467

Mangione, T. W., F. J. Fowler, and T. A. Louis (1992). Question characteristics and interviewer effects. Journal of Official Statistics, 8, 293–307.

McKenzie and Rosenzweig, (2012). Preface for symposium on measurement and survey design. Journal of Development Economics, vol. 98, 1-2

O'Muircheartaigh, C. and Campianelli, P. (1998). The relative impact of interviewer effects and sample design on survey precision. Journal of the Royal Statistical Society: Series A 161(1): 63-77.

Randall, S., Coast, E., Compaore, N. and Antoine, P. (2013). The power of the interviewer: a qualitative perspective on African survey data collection. Demographic research, 28 (27). 763-792.

Schaeffer, N. C. (1980). Evaluating race-of-interviewer effects in a national survey. Sociological Methods and Research. 8: 400-419.

Serneels, P. M., Beegle, K. G., and Dillon, A. S. (2016). Do returns to education depend on how and who you ask? Policy Research Working Paper Series 7747, World Bank.

Tripp, A.M, (2010). Museveni's Uganda: Paradoxes of power in a hybrid regime. Rienner.

West, B.T. and Olson, K. (2010). How much of interviewer variance is really nonresponse error variance. Public Opinions Quarterly, 74, 1004-1026

West, B.T., and Blom, A.G. (2016). Explaining interviewer effect: A research synthesis. Journal of Survey Statistics and Methodology, forthcoming

Table 1: Summary statistics

Variable	N	Mean	Std. Dev.
Enumerator			
Male	6895	0.353	0.478
High education	6693	0.922	0.269
Birth year	6895	1988	3
Urban	6895	0.610	0.488
Respondent			
Age	6895	44	14
Gender	6894	0.490	0.500
Single	6893	0.193	0.395
Years education	6894	5.814	4.689
Household size	6894	8.167	3.461
Asset index	6774	3.258	2.416
Number of cattle	6889	2.313	5.561
Consumes alcohol or tobacco	6889	0.413	0.492
Spending on alcohol or tobacco	6799	1543	5437
Spending on alcohol or tobacco, conditional	2756	3806	8021
Support for opposition party (DP)	6294	1.868	1.282
Support for opposition party (FDC)	6404	1.966	1.308
Support for national party (NRM)	6753	3.811	0.640
Support for opposition party (UPC)	6389	1.905	1.290

Notes: Table reports summary statistics for the full sample.

Table 2: Enumerator fixed effects and enumerator characteristics

	(1)	(2)	(3)	(4)
	Fixed effects	Enumerator characteristics	Enumerator characteristics + ability	Enumerator characteristics + ability + personality
Number of cattle	0.027	0.006	0.013	0.027
Single	0.004	0.001	0.002	0.004
Gender	0.009	0.004	0.006	0.009
Age	0.011	0.002	0.007	0.011
Household size	0.027	0.010	0.016	0.028
Amount spend on alcohol or tobacco	0.010	0.007	0.009	0.010
Consume alcohol or tobacco	0.009	0.003	0.004	0.009
Amount spend on alcohol or tobacco. conditional	0.023	0.013	0.019	0.023
Asset index	0.025	0.013	0.024	0.025
Education level	0.012	0.006	0.010	0.012
Support for DP	0.292	0.106	0.217	0.292
Support for FDC	0.234	0.089	0.162	0.234
Support for NRM	0.039	0.007	0.028	0.039
Support for UPC	0.257	0.087	0.184	0.258

Notes: Each column reports the sample average R^2. Results by region for each column/specification are reported in Appendix Tables A1-A4. Column (1) reports the adjuster R^2 from regressions that only include dummies for the enumerator fixed effect. Column (2) reports the adjusted R^2 from regressions that include observable enumerator characteristics (age, sex, education level and whether identify as urban or rural). Column (3) reports the adjusted R^2 from regressions that include the same enumerator characteristics as in column (2), plus a set of proxies for work ability (months of experience as enumerator, how important is that people think she/he is good at her/his work, whether she/he is strongly motivated by the money he can earn working as enumerator). Column (4) reports the adjusted R^2 from regressions that include the same enumerator characteristics as in column (3), plus additional personality characteristics (how much curiosity is the driving force behind what she/he does, how much she/he enjoys handling problems that are completely new, how much she/he enjoys trying to solve complex problems).

Table 3: Voting and enumerator characteristics

	(1) DP	(2) FDC	(3) NRM	(4) UPC	(5) DP	(6) FDC	(7) NRM	(8) UPC
Enumerator gender (male)	-0.557***	-0.399**	-0.114***	-0.512***	-0.362**	-0.357**	-0.047*	-0.398***
	(0.149)	(0.160)	(0.034	(0.116)	(0.158)	(0.142)	(0.026)	(0.131)
Enumerator education	0.010	-0.020	-0.0749*	-0.022	0.131	0.017	0.009	0.117
	(0.118)	(0.116)	(0.043)	(0.122)	(0.153)	(0.135)	(0.035)	(0.170)
Enumerator age	-0.022	-0.006	0.003	-0.034**	-0.008	-0.002	0.005	-0.029*
	(0.022)	(0.021)	(0.005)	(0.016)	(0.022)	(0.019)	(0.004)	(0.015)
Enumerator from urban locality	0.438**	0.543**	0.000	0.414***	0.647***	0.675***	0.006	0.568***
	(0.188)	(0.208)	(0.040)	(0.151)	(0.155)	(0.141)	(0.030)	(0.126)
Work experience as enumerator					0.006***	0.005***	0.000	0.004***
					(0.002)	(0.002)	(0.000)	(0.001)
Want people to know good can be at work					0.043	0.135	-0.027	0.043
					(0.141)	(0.127)	(0.020)	(0.114)
Strongly motivated by the wage can earn					-0.041	-0.096	0.048***	-0.075
					(0.093)	(0.088)	(0.013)	(0.083)
Enjoy handling new problems					0.054	0.028	-0.041	-0.053
					(0.148)	(0.133)	(0.025)	(0.138)
Enjoy trying to solve complex problems					0.126	0.073	0.078***	0.155**
					(0.082)	(0.068)	(0.018	(0.074)
Curiosity is a driving force for her actions					0.049	0.111	0.003	0.094
					(0.088)	(0.078)	(0.009)	(0.079)
Observations	6,245	6,333	6,553	6,320	5,947	6,033	6,251	6,018
R-squared	0.485	0.458	0.233	0.478	0.504	0.481	0.244	0.498

Note: This table presents OLS regression results for equation (1). For each column, the outcome variable is a measure of the stated support for a political party, namely DP, FDC, NRM, and UPC. Each regression includes community fixed effects and a constant. Robust standard errors in parenthesis are clustered at the enumerator level. *** p<0.01, ** p<0.05, * p<0.10.

Table 4: Alternative questions and enumerator characteristics for placebo test

	(1) Number of cattle	(2) Age	(3) Household size	(4) Alcohol and tobacco consumption
Enumerator gender (male)	0.115	0.241	0.094	0.017**
	(0.100)	(0.523)	(0.170)	(0.008)
Enumerator education	0.029	0.919	0.330	0.020
	(0.454)	-1.171	(0.297)	(0.016)
Enumerator age	-0.005	-0.047	0.013	0.003**
	(0.017)	(0.055)	(0.037)	(0.002)
Enumerator from urban locality	-0.216	0.340	0.221	-0.001
	(0.182)	(0.506)	(0.204)	(0.015)
Work experience as enumerator	0.002	-0.001	0.002	0.000
	(0.002)	(0.005)	(0.002)	(0.000)
Observations	5,947	6,033	6,251	6,018
R-squared	0.504	0.481	0.244	0.498

Note: This table presents OLS regression results for equation (1). For each column, the outcome variable is indicated in the first row. Additional controls not shown are as in Table 3. Each regression includes community fixed effects and a constant. Robust standard errors in parenthesis are clustered at the enumerator level. *** p<0.01, ** p<0.05, * p<0.10.

ONLINE APPENDIX

Table A1: Enumerator fixed effects

	Sample Average	**Team 1**	**Team 2**	**Team 3**	**Team 4**
Number of cattle	0.027	0.006	0.004	0.082	0.016
Single	0.004	0.005	-0.001	0.008	0.004
Gender	0.009	0.006	0.009	0.013	0.008
Age	0.011	0.006	0.010	0.012	0.021
Household size	0.027	0.013	0.019	0.044	0.039
Amount spend on alcohol or tobacco	0.010	0.005	0.008	0.007	0.026
Consume alcohol or tobacco	0.009	0.005	0.003	0.013	0.016
Amount spend on alcohol or tobacco. conditional	0.023	0.017	0.021	0.020	0.039
Asset index	0.025	0.029	0.010	0.035	0.025
Education level	0.012	0.003	0.014	0.022	0.011
Support for DP	0.292	0.434	0.186	0.166	0.346
Support for FDC	0.234	0.366	0.137	0.158	0.231
Support for NRM	0.039	0.057	0.012	0.063	0.012
Support for UPC	0.257	0.364	0.170	0.175	0.294

Notes: Table reports the R^2 from regressions that only include dummies for enumerator fixed effects.

Table A2: Observable enumerator characteristics

	Sample Average	Team 1	Team 2	Team 3	Team 4
Number of cattle	0.006	0.000	0.005	0.021	-0.001
Single	0.001	0.004	0.000	-0.001	0.000
Gender	0.004	0.006	0.000	0.006	0.004
Age	0.002	-0.001	0.009	0.000	0.003
Household size	0.010	0.005	0.015	0.010	0.012
Amount spend on alcohol or tobacco	0.007	0.003	0.004	-0.001	0.026
Consume alcohol or tobacco	0.003	0.003	0.003	0.000	0.008
Amount spend on alcohol or tobacco. conditional	0.013	-0.004	0.016	0.011	0.039
Asset index	0.013	0.007	0.003	0.030	0.013
Education level	0.006	0.002	0.010	0.008	0.004
Support for DP	0.106	0.173	0.089	0.009	0.141
Support for FDC	0.089	0.161	0.055	0.018	0.102
Support for NRM	0.007	0.009	0.007	0.005	0.007
Support for UPC	0.087	0.138	0.076	0.007	0.118

Notes: Table reports the adjusted R^2 from regressions of the dependent (row) variable on observable enumerator characteristics (age, sex, education level, and whether identify as urban or rural).

Table A3: Enumerator characteristics and ability

	Sample Average	Team 1	Team 2	Team 3	Team 4
Number of cattle	0.013	0.003	0.005	0.039	0.007
Single	0.002	0.004	-0.001	0.003	0.002
Gender	0.006	0.007	0.007	0.007	0.004
Age	0.007	0.006	0.01	0.001	0.015
Household size	0.016	0.009	0.021	0.01	0.031
Amount spend on alcohol or tobacco	0.009	0.007	0.006	0.001	0.027
Consume alcohol or tobacco	0.004	0.005	0.002	0.001	0.01
Amount spend on alcohol or tobacco. conditional	0.019	0.017	0.017	0.008	0.04
Asset index	0.024	0.028	0.007	0.035	0.025
Education level	0.010	0.005	0.01	0.019	0.007
Support for DP	0.217	0.275	0.168	0.099	0.328
Support for FDC	0.162	0.246	0.114	0.065	0.208
Support for NRM	0.028	0.031	0.012	0.052	0.014
Support for UPC	0.184	0.247	0.138	0.081	0.266

Notes: Table reports the adjusted R^2 from regressions of the dependent (row) variable on observable enumerator characteristics (age, sex, education level; whether she/he identifies as urban or rural) and a set of proxies for worker's ability (months of experience as enumerator; how important is that people think he is good at his/her work; whether she/he is strongly motivated by the money she/he can earn working as enumerator).

Table A4: Enumerator characteristics with (selected) personality measures

	Sample average	Region 1	Region 2	Region 3	Region 4
Number of cattle	0.027	0.006	0.004	0.082	0.016
Single	0.004	0.005	-0.001	0.008	0.004
Gender	0.009	0.006	0.009	0.013	0.008
Age	0.011	0.006	0.010	0.012	0.021
Household size	0.028	0.013	0.021	0.044	0.039
Amount spend on alcohol or tobacco	0.010	0.005	0.008	0.007	0.026
Consume alcohol or tobacco	0.009	0.005	0.003	0.013	0.015
Amount spend on alcohol or tobacco, conditional	0.023	0.017	0.021	0.02	0.039
Asset index	0.025	0.029	0.01	0.035	0.025
Education level	0.012	0.003	0.014	0.022	0.011
Support for DP	0.292	0.434	0.188	0.166	0.346
Support for FDC	0.234	0.366	0.14	0.158	0.231
Support for NRM	0.039	0.057	0.013	0.063	0.012
Support for UPC	0.258	0.364	0.172	0.175	0.294

Notes: Table reports the adjusted R^2 from regressions of the dependent (row) variable on observable enumerator characteristics (age, sex, education level, and whether she/he identifies as urban or rural), a set of proxies for worker's ability (months of experience as enumerator, how important is that people think he is good at his/her work, and whether she/he is strongly motivated by the money he can earn working as enumerator), and proxies for personality characteristics (how much curiosity is the driving force behind what she/he does; how much she/he enjoys handling problems that are completely new; how much she/he enjoys trying to solve complex problems).

Table A5 - Robustness: Voting and enumerator characteristics (different enumerator sample)

	(1) DP	(2) FDC	(3) NRM	(4) UPC	(5) DP	(6) FDC	(7) NRM	(8) UPC
Enumerator gender (male)	-0.515***	-0.302**	-0.091***	-0.444***	-0.387***	-0.234*	-0.044*	-0.363***
	(0.121)	(0.140)	(0.028)	(0.101)	(0.094)	(0.118)	(0.022)	(0.095)
Enumerator education	0.108	0.095	-0.094**	0.108	0.233*	0.130	-0.050	0.236
	(0.112)	(0.123)	(0.038)	(0.114)	(0.139)	(0.139)	(0.042)	(0.145)
Enumerator age	-0.033*	0.002	0.003	-0.040***	-0.029	0.008	0.002	-0.039***
	(0.017)	(0.019)	(0.003)	(0.013)	(0.020)	(0.020)	(0.003)	(0.012)
Enumerator from urban locality	0.439***	0.564***	-0.030	0.440***	0.576***	0.675***	-0.031	0.571***
	(0.162)	(0.184)	(0.038)	(0.135)	(0.136)	(0.129)	(0.035)	(0.114)
Work experience as enumerator					0.005***	0.006***	0.000	0.004***
					(0.001)	(0.002)	(0.000)	(0.001)
Want people to know good can be at work					0.056	0.099	-0.046**	0.067
					(0.092)	(0.091)	(0.019)	(0.078)
Strongly motivated by the wage can earn					0.008	-0.068	0.040***	-0.037
					(0.086)	(0.087)	(0.012)	(0.081)
Enjoy handling new problems					0.136	0.149	-0.003	0.039
					(0.129)	(0.111)	(0.030)	(0.107)
Enjoy trying to solve complex problems					0.091	0.028	0.069***	0.123*
					(0.081)	(0.071)	(0.019)	(0.073)
Curiosity is a driving force for her actions					-0.173*	-0.077	-0.030	-0.136
					(0.094)	(0.088)	(0.021)	(0.084)
Observations	6,761	6,868	7,175	6,854	6,401	6,504	6,807	6,488
R-squared	0.474	0.439	0.226	0.464	0.490	0.462	0.236	0.481

Note: This table presents OLS regression results for equation (1). For each column, the outcome variable is a measure of the stated support for a political party, namely DP, FDC, NRM, and UPC. Each regression includes community fixed effects and a constant. For all regression, the sample includes enumerators who have completed at least 24 interviews. Robust standard errors in parenthesis are clustered at the enumerator level. *** p<0.01, ** p<0.05, * p<0.10.

Table A6 - Robustness: Including respondent characteristics as controls

	(1) DP	(2) FDC	(3) NRM	(4) UPC	(5) DP	(6) FDC	(7) NRM	(8) UPC
Enumerator gender (male)	-0.554***	-0.401**	-0.112***	-0.509***	-0.359**	-0.360**	-0.045	-0.396***
	(0.150)	(0.159)	(0.035)	(0.115)	(0.158)	(0.142)	(0.027)	(0.130)
Enumerator education	0.015	-0.018	-0.074*	-0.0199	0.139	0.0246	0.008	0.121
	(0.121)	(0.117)	(0.043)	(0.124)	(0.156)	(0.137)	(0.036)	(0.171)
Enumerator age	-0.022	-0.006	0.003	-0.033*	-0.008	-0.002	0.005	-0.028*
	(0.022)	(0.021)	(0.005)	(0.016)	(0.023)	(0.020)	(0.004)	(0.015)
Enumerator from urban locality	0.440**	0.543**	0.000	0.414***	0.648***	0.667***	0.009	0.567***
	(0.187)	(0.205)	(0.040)	(0.150)	(0.156)	(0.141)	(0.032)	(0.125)
Work experience as enumerator					0.006***	0.005***	0.001	0.004***
					(0.002)	(0.002)	(0.000)	(0.001)
Want people to know good can be at work					0.043	0.133	-0.026	0.0436
					(0.142)	(0.129)	(0.020)	(0.115)
Strongly motivated by the wage can earn					-0.038	-0.089	0.046***	-0.073
					(0.093)	(0.087)	(0.013)	(0.082)
Enjoy handling new problems					0.051	0.024	-0.040	-0.055
					(0.149)	(0.135)	(0.025)	(0.139)
Enjoy trying to solve complex problems					0.121	0.065	0.079***	0.149*
					(0.083)	(0.069)	(0.017)	(0.074)
Curiosity is a driving force for her actions					0.049	0.111	0.002	0.094
					(0.089)	(0.079)	(0.001)	(0.079)
Respondent age	-0.000	-0.000	0.001	0.001	-0.001	-0.001	0.001	0.0004
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Respondent gender (male)	-0.007	0.066*	-0.040*	0.013	-0.011	0.069**	-0.039*	0.024
	(0.030)	(0.034)	(0.021)	(0.033)	(0.031)	(0.033)	(0.022)	(0.034)
Respondent marital status (single)	0.084	0.120**	-0.031	0.066	0.084	0.123**	-0.016	0.072
	(0.059)	(0.057)	(0.029)	(0.060)	(0.059)	(0.056)	(0.028)	(0.060)
Respondent education (writing)	-0.041*	-0.052**	0.007	-0.037*	-0.041*	-0.048**	0.006	-0.034
	(0.023)	(0.022)	(0.014)	(0.022)	(0.023)	(0.023)	(0.015)	(0.023)
Observations	6,243	6,330	6,550	6,317	5,946	6,031	6,249	6,016
R-squared	0.486	0.460	0.234	0.480	0.505	0.483	0.245	0.498

Note: This table reports OLS regression results for equation (1) and includes respondent characteristics. For each column, the outcome is a variable measuring the stated support for a political party, namely DP, FDC, NRM, and UPC. Each regression includes community fixed effects and a constant. Robust standard errors in parentheses are clustered at the enumerator level. *** p<0.01, ** p<0.05, * p<0.10.

Table A7 - Robustness: Using a dummy to measure political opinions

	(1) DP	(2) FDC	(3) NRM	(4) UPC	(5) DP	(6) FDC	(7) NRM	(8) UPC
Enumerator gender (male)	-0.186***	-0.117*	-0.022*	-0.178***	-0.110**	-0.113**	-0.014	-0.135***
	(0.053)	(0.060)	(0.012)	(0.042)	(0.052)	(0.052)	(0.012)	(0.044)
Enumerator education	-0.001	-0.006	-0.020*	-0.012	0.048	0.006	-0.013	0.040
	(0.038)	(0.042)	(0.012)	(0.039)	(0.044)	(0.043)	(0.014)	(0.050)
Enumerator age	-0.005	0.000	0.002	-0.012**	0.001	0.002	0.002	-0.010*
	(0.008)	(0.008)	(0.001)	(0.006)	(0.009)	(0.008)	(0.002)	(0.005)
Enumerator from urban locality	0.145**	0.206***	0.014	0.143***	0.221***	0.240***	0.005	0.202***
	(0.064)	(0.075)	(0.013)	(0.053)	(0.052)	(0.052)	(0.015)	(0.043)
Work experience as enumerator					0.002***	0.002***	-0.000	0.001***
					(0.001)	(0.001)	(0.000)	(0.000)
Want people to know good can be at work					0.018	0.046	-0.004	0.026
					(0.048)	(0.048)	(0.009)	(0.039)
Strongly motivated by the wage can earn					-0.008	-0.030	0.014*	-0.019
					(0.031)	(0.032)	(0.007)	(0.028)
Enjoy handling new problems					0.025	0.004	-0.001	-0.017
					(0.045)	(0.047)	(0.009)	(0.042)
Enjoy trying to solve complex problems					0.031	0.015	0.010*	0.046**
					(0.025)	(0.025)	(0.006)	(0.021)
Curiosity is a driving force for her actions					0.019	0.039	-0.002	0.029
					(0.030)	(0.029)	(0.005)	(0.026)
Observations	6,245	6,333	6,553	6,32	5,947	6,033	6,251	6,018
R-squared	0.468	0.436	0.222	0.451	0.488	0.461	0.228	0.470

Note: This table presents OLS regression results for equation (1). For each column, the outcome variable is dummy which takes value 1 if the respondent is open to vote for the political party indicated in the column (namely DP, FDC, NRM, and UPC) and zero otherwise. Each regression includes community fixed effects and a constant. Robust standard errors in parenthesis are clustered at the enumerator level. *** p<0.01, ** p<0.05, * p<0.10.

Table A8 - Robustness: Including respondent-enumerator interactions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	DP	FDC	NRM	UPC	DP	FDC	NRM	UPC
Respondent gender (male)	-0.045	0.013	-0.067***	-0.019	-0.041	0.007	-0.066***	-0.0130
	(0.037)	(0.043)	(0.021)	(0.041)	(0.036)	(0.042)	(0.022)	(0.0390)
Enumerator gender (male)	-0.609***	-0.477***	-0.149***	-0.556***	-0.406**	-0.454***	-0.084**	-0.453***
	(0.172)	(0.172)	(0.033)	(0.129)	(0.170)	(0.149)	(0.033)	(0.136)
Respondent gender * Enumerator gender	0.108	0.149	0.075*	0.093	0.0894	0.179*	0.078*	0.110
	(0.102)	(0.095)	(0.039)	(0.099)	(0.101)	(0.095)	(0.043)	(0.104)
Respondent age	-0.001	-0.000	0.001	0.001	-0.001	-0.001	0.001	0.000473
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.00113)
Respondent marital status (single)	0.086	0.122**	-0.030	0.068	0.085	0.125**	-0.015	0.0735
	(0.058)	(0.056)	(0.029)	(0.060)	(0.059)	(0.056)	(0.028)	(0.0597)
Respondent education	-0.042*	-0.053**	0.006	-0.038*	-0.042*	-0.049**	0.0053	-0.0356
	(0.023)	(0.022)	(0.014)	(0.022)	(0.023)	(0.023)	(0.013)	(0.0232)
Enumerator age	-0.022	-0.010	0.003	-0.033**	-0.008	-0.003	0.005	-0.0288*
	(0.022)	(0.021)	(0.005)	(0.016)	(0.023)	(0.019)	(0.004)	(0.0156)
Enumerator education	0.019	-0.013	-0.072	-0.017	0.141	0.028	0.010	0.123
	(0.121)	(0.118)	(0.043)	(0.125)	(0.156)	(0.138)	(0.036)	(0.172)
Enumerator from urban locality	0.440**	0.543**	0.000	0.414***	0.647***	0.665***	0.008	0.565***
	(0.187)	(0.206)	(0.039)	(0.151)	(0.156)	(0.141)	(0.032)	(0.125)
Work experience as enumerator					0.006***	0.005***	0.000	0.00405***
					(0.002)	(0.002)	(0.000)	(0.00145)
Want people to know good can be at work					0.044	0.133	-0.026	0.0436
					(0.141)	(0.128)	(0.020)	(0.114)
Strongly motivated by the wage can earn					-0.039	-0.091	0.045***	-0.0738
					(0.093)	(0.087)	(0.013)	(0.0822)
Enjoy handling new problems					0.051	0.025	-0.039	-0.0543
					(0.149)	(0.134)	(0.026)	(0.139)
Enjoy trying to solve complex problems					0.122	0.067	0.079***	0.151**
					(0.083)	(0.069)	(0.018)	(0.0745)
Curiosity is a driving force for her actions					0.0491	0.113	0.003	0.0951
					(0.089)	(0.078)	(0.010)	(0.0794)
Observations	6,243	6,330	6,550	6,317	5,946	6,031	6,249	6,016
R-squared	0.486	0.461	0.235	0.480	0.505	0.484	0.245	0.499

Notes: This table reports OLS regression results for equation (1) and includes respondent characteristics and an interaction with enumerator gender. For each column, the outcome variable is variable measuring the stated support for a political party, namely DP, FDC, NRM, and UPC. Each regression includes community fixed effects and a constant. Robust standard errors in parentheses are clustered at the enumerator level. *** p<0.01, ** p<0.05, * p<0.10.