

Abstract

There is an inherent tension between traditional norms between interviewer and respondent characteristics. More and survey protocols for quantitative data collected in the objective measures serve as a pseudo control group. In the developing world. Unexpected interactions between the paper finds interviewer effects in the both subjective and interviewer and respondent can lead to interviewer effects in objective data, but the magnitude is considerably stronger in the data, particularly in the case of subjective or sensitive subjective questions. The paper also finds that female questions. This paper makes use of a unique data set available respondents are more susceptible to influence based on able from Timor-Leste containing subjective and objective interviewer's beliefs. Despite methodological short questions to study these effects. In addition to their age groupings, the study highlights the need to consider more and gender, data were collected from the interviewers by the impact of traditional cultural norms when conducting their opinions on the subjective questions prior to conducting quantitative surveys in the developing world on to fieldwork. Fixed effects and mixed effects logit models to topics that are outside the standard objective questions. are used to examine the main effects and interactions

This paper is a product of the Poverty Global Practice Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The author may be contacted at khimelein@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Interviewer Effects in Subjective Survey Questions:
Evidence from Timor-Leste

Kristen Himelein¹

World Bank

1818 H Street NW, Washington DC, 20433, USA

Key words: survey methodology, interviewer effects, subjective data, justice surveys, Timor-Leste

¹ Kristen Himelein is a Senior Economist / Statistician in the World Bank's Poverty Global Practice, and can be reached at khimelein@worldbank.org. The author would like to thank Kinnon Scott, Steve Heeringa, Brady West, Melissa Plegue, the participants in the 2013 European Survey Research Association conference in Ljubljana, and two anonymous referees for very useful comments. This analysis was supported by Research Support Budget program in the Development Economics Research Group of the World Bank. All views are those of the author and do not reflect the views of the World Bank or its member countries.

1. Introduction

Collecting data through quantitative interviews can be an awkward social interaction. An inherent tension is created between the societal norms of conversation and the survey protocols needed to standardize information across a range of interviewers and respondents (Kahn & Cannell, 1957, Suchman & Jordon, 1980). Individual interviewers may have differing levels of success in navigating this interaction, creating interviewer effects in the data. The further the interview is from normal conversation, such as discussing sensitive topics with a stranger, the more likely it is to find non-trivial impacts from the interviewer (Schnell & Kreuter, 2005).

The tension between culture and protocol is often amplified in developing country settings. Traditional societies often have strict cultural codes of behavior. Interviewing is also generally conducted face-to-face, and therefore more reminiscent of a social setting than a telephone or e-mail survey. In Timor-Leste, traditional society is particularly hierarchical, with a history of exclusion of women, minorities, and the youth. In addition data collection for this study took place in 2008, following the end of decades of conflict. While political violence and other organized aggression was rare at this time, simmering ethnic and regional tensions remained. The questions in this study focus on the respondent's perceptions of corruption, community dispute resolution, women's rights, and land claims, all potentially contentious topics.

Two main methodological questions are addressed in this paper. First, are interviewer effects present in the survey data, and, if so, is the magnitude greater for the sensitive subjective questions compared to more objective questions? Secondly, what, if any conclusions can be drawn regarding the nature of these interviewer effects? The main hypothesis is that if interviewer effects were present, respondents would likely show a form of social desirability bias (Crowne and Marlowe, 1960), deferring to the traditional norms of hospitality to give the answer they perceive the interviewer "guest" would like to hear. The respondent's expectations are based on the interviewer's observable traits and possible additional cues given by the interviewer based on their opinions. The expectation is that this would lead to pro-state responses to young or female interviewers and more conservative responses to older male interviewers.

The analysis makes use of a unique data set available from Timor-Leste which specifically examines the characteristics of the respondent, the characteristics of the interviewer, and the interviewer's personal opinions on the survey questions. It also briefly addresses the question of salience with a more detailed analysis of gender and the women's rights question. The analysis uses fixed effects logit and multi-level mixed-effects logistic regression modeling to test for the interviewer effects. Once the presence of these effects has been identified, I attempt to estimate the amount of total variation they are capable of explaining, and to estimate the explanatory power of the three available interviewer characteristics, age, gender, and the interviewer's opinion, as well as the associated interaction terms.

The next section briefly reviews existing literature on interviewer effects as it relates to the research questions. Section 3 discusses the data, results are presented in section 4, and section 5 offers a brief conclusion and possible areas of future research.

2. Literature Review

The question of the interviewers' effects on the respondent's answers has been addressed extensively in the social science literature, and the findings vary widely depending on the variable under discussion and the context of the survey. A common mechanism by which interviewers can introduce bias is through an action on the part of the interviewer that deviates from survey protocols, either intentional or unintentional, verbal or non-verbal. Though training is designed to prevent this behavior, interviewers often deviate (Billiet & Loosevelt, 1988; Fowler & Mangione, 1990). In particular, Cannell, Miller & Oksenberg (1981) note a tendency to elaborate and paraphrase difficult questions. The questions in the Timor-Leste study would be particularly vulnerable to this as they use an unfamiliar forced-choice format and address slightly abstract topics. These questions were also both subjective and sensitive, which the literature has found to be more vulnerable to interviewer effects, though there are some exceptions. See Cotter, Cohen & Coulter (1982), Davis (1997), Davis, Couper, Janz, Cadwell & Resnicow (2010), Groves (1989), Samples et al (2014), Schaeffer (1980), and Schnell & Kreuter (2005) for further discussion.

In addition to explicit actions, it is also possible to introduce bias even while following protocols if the respondent alters their answer based on the physical characteristics of the interviewer, such as age, race, and gender. While findings again vary with context, most literature find significant impacts, particularly with race and gender. See Collins & Butler (1983), Davis (1997), Ellison, McFarland & Krause (2011), Finkel, Geterbock & Borg (1991), Freeman & Butler (1976), Galla, Frisone, Jeffrey & Gaer (1981), Hatchett & Schuman (1975), Krysan & Cooper (2003), Rhodes (1994), Singer, Frankel & Glassman (1983) for examples. In addition, these effects may be amplified for sensitive issues, such as gender and sexual practices (Tu & Liao, 2007; Johnson & Moore, 1993; Lueptow, Moser & Pendleton, 1990; Johnson & DeLamater, 1976; Galla, et al, 1981). The results from the literature are more mixed with regard to interaction effects. For example, while Huddy et al (1997) found an average increase of five percentage points on a dichotomous question for female-to-female interviews, and similar findings are recorded by Lueptow et al (1990) and Landis, Sullivan & Sheley (1973), mixed results are found by Ballou (1990), Johnson & DeLamater (1976) and Catania et al (1996). Related to age, Wallace, Kohout & Colsher (1992) found that older respondents were more comfortable with older interviewers.

The literature is more limited on the impact of the interviewer's beliefs on the respondent's answers. In one of the earliest examples of interviewer effects in the literature, Rice (1929) found that an interviewer's personal politics, in this case socialism versus prohibition, led to wide differences the main cause of destitution cited by respondents. More recently Hermann (1983, discussed in Schnell & Kreuter, 2005) found "interviewer effects were small if the item was unimportant to the interviewer and that they were large if the item was unimportant to the respondent."

There are fewer examples of studies conducted outside of the United States. Phung, Hardeweg, Praneetvatakul & Waibel (2013) found significantly higher errors based on interviewer

gender, ethnicity, age, and education in a study of Thailand and Vietnam, and concluded based on interaction effects that matching on respondent and interviewer characteristics improved data quality. Liu & Stainbeck (2013) found inconsistent gender of interviewer effects in a survey on marriage conducted in China. Blaydes & Gillum (2013) examined the impact of a veiled enumerator on questions of religiosity in Egypt, and found Muslim women indicated they were more religious to a veiled enumerator while Christians indicated less. These effects were the largest for younger, poorest and less educated women. Benstead (2010) conducted a similar survey in Morocco that included both male and female respondents and interviewers, finding that respondents gave more socially progressive responses to female interviewers, though less so for veiled female interviewers. She also finds that women respondents were more likely to give socially progressive responses to female interviewers. Flores-Macias & Lawson (2008) find that respondents in Mexico City gave more socially progressive answers when questioned by a woman, but found no differences in the remaining districts of the country. Meng-Li & Yu (2008) looked at the effect education differentials on biotechnology surveys in Taiwan, China, and found no significant first order effects on age, education or interviewer experience, but did find significance on certain second order interactions related to education.

There are other aspects of the interviewer effects literature that are not relevant to this discussion. For example much of the recent literature has focused on the impact of interviewer-correlated non-response on bias and the precision of estimates (see West & Olson, 2010 for a recent example). In the case of the Timor-Leste survey, agreement to participate was secured jointly by the team supervisor and village officials at the cluster level, and therefore it would not be possible to separate the supervisor and geography effects for unit non-response. Unit non-response is also less than five percent. For item non-response, all respondents answered all five questions considered here, though some indicated “both” or “neither” of the two choices offered. (See Tables 2 and 3 in the appendix for question-level response rates.) These options were included to decrease the probability that the interviewer would rephrase the question in order to elicit an answer from a potentially confused respondent, and to not overly pressure a respondent to answer a potentially sensitive item. This analysis is not considered in this paper but may be for future work. Other areas of the literature which are not applicable here include learning, as the survey is a cross section (see Bollinger & David, 2005; Duan et al, 2007; and Olson & Peytchev, 2007), and incentive payments as interviewers are paid a flat salary for participation (Kennickell, 2004; Rosen et al, 2010).

3. Data

In 2007 the Timor-Leste Survey of Living Standards (TLSLS2) collected information on household demographics, consumption based welfare measures, and a range of other socioeconomic characteristics. The following year, the World Bank conducted an extension survey (TLSLS-2X) to collect additional information on agriculture, shocks and vulnerability,

financial services, and access to justice. The extension survey, which is the main data set in this study, revisited a representative subsample of 1,427 households from the original TLSLS2, though was not a panel survey as different questions were asked during the revisit.

Fieldwork was conducted between April and October 2008 by seven mobile field teams of three interviewers and one supervisor. Interviewer teams were not constructed based on specific criteria, though efforts were made for each team to contain at least one woman. Teams were then assigned to constructed geographic areas of the country called zones. (Zones, generally though not always, corresponded to groups of administrative districts.) Within teams, interviewers were randomly assigned to households.² Supervision missions were taken to monitor the quality of fieldwork. A subsequent review of data quality led to the exclusion of data from one field team, and the re-sampling and re-interview of this zone by the remaining six teams. Additionally, during the course of the survey, there were replacements made for three interviewers – for which interviewer-level information was not available. These observations are therefore excluded, leaving a total of 19 interviewers available for analysis.

The original TLSLS2 sample was a two-stage probability sample, with selection of PSUs proportional to size in the first stage and simple random sampling of households in the second stage. The extension sample was selected as a nationally representative simple random sample of PSUs. In selected PSUs, 12 of the 15 original households were randomly selected to be revisited, with the final three households being held as replacements. Overall, the survey had a seven percent replacement rate. The respondent for the justice module was randomly selected to be either the head of household or spouse of the household head to ensure an equal gender distribution of men and women in the sample. Sampling weights were calculated as the inverse probability of overall selection, with a post-estimation correction to the stratum population totals.

The access to justice module was specifically designed to capture indicators that are traditionally difficult to measure quantitatively. The questions used a similar structure to subjective questions found in the Afrobarometer and Latinobarometer surveys, which offer respondents two neutral statements and ask which is closer to the respondent's opinion. The full text of the questions is available in the appendix but generally the first question relates to corruption (Corruption), the second to whether the main responsibility for law and order lies with the police or the community (Law & Order), the third to a woman's role in traditional society

² This was in accordance with the field directions to even distribute workload. Since households in Timor-Leste can be quite distant and sometimes only accessible by foot, randomly assigning the interviewers within a PSU is necessary for an even workload. The assumption was also tested, using t-tests, and taking into account the complex sample design when calculating the standard errors. There are no statistically significant differences in the average age, gender, or years of education of the household head, or the average household consumption, between male and female interviewers, or between interviewers older or younger than the median interviewer age of 33.

(Women's Rights), the fourth to the resolution of non-violent disputes (Dispute Resolution), and finally to land titling and conflict in communities (Land Titling).³

Fieldworkers for the project were experienced enumerators, most of whom had participated in previous data collection efforts on behalf of the statistics bureau. They underwent a rigorous week-long training, including numerous field sessions. After the training but prior to the start of fieldwork, interviewers and supervisors were asked to complete the perceptions of the law section based on their own beliefs. Interviewers were not told they were participating in a methodological study; rather that it was an evaluation exercise as part of the training.

In general, the characteristics of the interviewers differed from those in the general population. Interviewers were younger and more likely to be male than respondents, and more highly educated as interviewers were required to have a minimum of completed secondary education, while only 12 percent of respondents met this criterion. (See Table 1 for further details.) The interviewer's responses to the subjective questions were generally consistent with those given by respondents with similar characteristics. Limiting the respondent sample to urban residents with secondary education, and controlling for age and gender, there are statistically significant differences between interviewers and respondents only for question 4, Dispute Resolution. This is likely related to the fact that all of the interviewers live in the capital, Dili, were residents were in general less likely to indicate a preference for community over formal dispute resolution systems, though the limited sample size precludes a more formal test. Table 2 shows the comparison between the interviewer's responses and those of the overall population.

As a pseudo comparison group for the subjective questions, I also test four more objective dichotomous variables also collected in the extension survey: whether or not the respondent attended a community meeting in the last 12 months, if the household had experienced problems with Siam Weed (an invasive species common to rural Timor-Leste), if the household's farm had stone terraces (a wide spread agricultural practice), and whether the household had experienced a drought in the last two years. The full text of these questions is listed in the appendix. Table 3 shows the distribution and non-response rates for the objective variables.

4. Results

4.1 Fixed Effects Model

³ To simplify the presentation of results, the five questions are given names linked to the general issue to which they relate. These names are meant as a shorthand description only. The actual phrasing of the question was designed to be neutral, and the full text is available in the appendix.

The first step in the analysis was to identify if interviewer effects were present in the data, and, if so, to what degree. Table 4 shows an excerpt of the results for a fixed effects logit model for the five subjective measures described above, as well as four more objective measures. The dependent variable for the subjective measures is whether the respondent said A, or that the first statement was closed to their beliefs. For the objective questions, the dependent variable is a ‘yes’ response. In addition to the interviewer fixed effects, the full model includes the gender, age, education level of the respondent, urban/rural residence, the log per capita expenditure of the household, and the district of residence. The interviewers are identified by their gender and age. Both sets of models take into account the complex design features, stratification and clustering of the survey.

The results indicate strong jointly significant for interviewer effects in all nine models. Though the interviewer effects are significant in both the subjective and objective models, they explain a larger percentage of the total variation in the subjective models. Table 5 shows the Nagelkerke adjusted R^2 values for identical logit models with and without the interviewer fixed effects.⁴ The percentage of the variation explained by the interviewer fixed effects ranges between 41.0 percent and 46.1 percent for the five subjective questions, and between 11.8 percent and 25.7 percent in the four objective models. These percentages, however, likely overstate the percentage of the variation attributable solely to interviewer effects. Because of the field team structure and grouped geographic assignment of the teams, it is not possible to fully separate the interviewer and geographic variation, even with the inclusion of district fixed effects in the model. These impacts could be substantial as post-conflict Timor-Leste was characterized by high levels of fractionalization and regional differences. This issue though would also occur for the objective measures, and therefore the comparisons are likely robust. These findings support many previous studies in the literature on interviewer effects in sensitive and subjective questions.

4.2 Mixed Effects Model

Having established the presence of interviewer effects, a mixed effect logit is used to allow for the inclusion of interviewer-level covariates in addition to interviewer random effects. For this analysis, it is not possible to include the weights nor adjust the standard errors for clustering or stratification when implemented using Stata software for statistical analysis. While the design is approximately self-weighting within the stratum, the exclusion of the clustering would underestimate standard errors.

The specification for the model is, for each of the five questions:

⁴ Note that these models do not use the complex design features of the data because the ratios of likelihoods with complex sample design are not statistically valid.

$$\Pr(Y_j = 'A') = \frac{1}{1+e^{-\eta}} \text{ where } \eta = \alpha + \beta_1 X_j + \beta_2 W_i + \beta_3 I_i + \mu_i + \varepsilon_{ij}$$

The dependent variable Y_j is a binary variable for a given question response ('A' in this case) for respondent j , α is a constant term, X is a vector of relevant household and individual characteristics for respondent j , W is a set of characteristics for interviewer i , I is the interviewer's personal response to an individual question, μ_j is the random effect associated with the intercept for interviewer i with a distribution of

$$\mu_i \sim N(0, \sigma_i^2)$$

where σ_i^2 represents the variance of the random interviewer effects. The residual error for respondent j with interviewer i for each question is represented by ε_{ij} with a distribution of the error term of

$$\varepsilon_{ij} \sim N(0, \sigma^2)$$

The variables included in vector X are the respondent's gender, age, and level of education, and the household's log per capita household expenditure, urban/rural residence status, and district of residence. The interviewer level characteristics include age and gender only, as the minimum level of education does not vary across interviewers. Results of the mixed effects model are presented in Table 6.⁵

While the individual respondent-level predictors for the question response are not the main focus of this analysis, but it is important to note, as demonstrated in Table 5, some are significant and collectively they explain the majority of the variation for a given question. For the Corruption question, urban residents are less likely to sanction 'additional payments' to public officials, while women and wealthier respondents are more likely to condone such payments. There are no statistically significant effects on the respondent's age or education. From the coefficients on the education variables there seemed to be a trend in the variables of higher educated people be more likely find additional payments permissible, but even tested jointly the results are not significant. In the Law & Order question, wealthier respondents are more likely to indicate that the main responsibility for maintaining law and order within the community lies with the police, as opposed to traditional leaders and the community itself. Results from the Women's Rights question indicate, unsurprisingly, that women are more likely to believe that a woman should be allowed to speak for herself in the traditional *adat* dispute resolution proceedings. There are no statistically significant effects with regard to age, education, urban/rural residence status, or household

⁵ Due to the nature of the data, multicollinearity may also be an issue. Multicollinearity results when there is a correlation between independent variables in regression analysis. In the case of severe multicollinearity, the variance is overestimated, thereby causing the false acceptance of the null hypothesis. In addition, there can be incorrect signs on the coefficients for explanatory variables. To test for multicollinearity, the models including the interviewer's gender, age, and opinion on the subjective questions were re-run as linear regressions following which variance inflation factors (VIF) values were calculated. The VIF were all below the threshold of concern of five, with nearly all being also below three.

consumption. For the Dispute Resolution question and the Land Titling question, with the exception of some district dummy variables, none of the individual characteristic explanatory variables are statistically significant. Across the five subjective models, the geographical indicator variables for Timor-Leste's 13 districts are in many cases significant.

The results of this model also support the finding of interview effects from the fixed effects logit model. In the mixed effects model, the random intercepts are significant for all five subjective models, indicating that the individual interviewers do impact the outcome. This is more general than fixed effects since the fixed effects model calculates only the presence of interview effects for the specific set of interviewers tested, while the mixed effects model assumes that there is a universe of possible interviews, of which a random sample is selected for participation in the survey.

To examine potential outlier effects of individual interviewers, I calculate the empirical best linear unbiased predictor for each interviewer on each of the five subjective questions and test for outlier values. Of the 93 interviewer-question combinations that were estimated by the models, five are significant at the ten percent or greater threshold of significance, and two at the five percent threshold. Those which are significant are not centered on a particular interviewer or question, and are generally consistent with what would be expected to find in terms of frequency of outliers in a normal distribution. Following Wilk and Gnanadesikan (1968), graphs comparing quantiles of empirical best linear unbiased predictors against quantiles of normal distribution are shown in Figure 1 in the appendix. These graphs show the effects generally fall on or near the 45 degree line, though there are single outlier values for Corruption and Dispute Resolution.

There are limited significant results for the first order effects on age, gender and the interviewer's beliefs, even though the exclusion of the weights and complex design features lead to the underestimation of standard errors. Respondents were more likely to tell a female respondent that the main responsibility for maintaining law and order in the community was with the police instead of the traditional *chefe de suco*. This is consistent with respondents giving more pro-state answers to historically disenfranchised groups. If this hypothesis were true, however, one would expect positive and significant results on the Women's Rights and Land Titling questions, and negative and significant results on Corruption and Dispute Resolution. None of the other questions show significance, and the signs on Corruption and Land Titling move in the opposite direction. With regard to age, assuming the youth to be more progressive, the expected results would be the same as above for female interviewers. All but the Women's Rights question show the expected sign, but only Land Titling has a significant result. For older interviewers, which assumedly would be perceived as more conservative, the expectation on the signs would be in the opposite direction. Three of the five questions have the expected sign, but only one, Law & Order, has a significant result. Overall, the first order results show general support for the hypothesis, but the results are far from conclusive.

In all five questions, the respondent is more likely to agree with choice 'A' if the interviewer also selected 'A' though this difference is weakly significant in two cases (at the 10% level), and

not significant in the remaining three cases. The coefficients on the two weakly significant cases indicate substantial effects however. Respondents are almost twice as likely to respond ‘A’ when that was the response of the interviewer for the Corruption and Law & Order questions.

4.3 Comparisons to Objective Variables

As a verification check that the results of tests performed are not anomalies common across the data set and unrelated to interviewer effects, the mixed effects models are also calculated for the four objective variables. The main results follow expectations for Timor-Leste. For example older, while rural and less educated farmers being less likely to have stone terraces on their land. Again, many of the district level effects are significant. With regard to the interviewer-level effects, the only effect significant at the 95% confidence level is that respondents are less likely to tell interviewers under 30 years of age that they have experienced drought. There is also a weakly significant relationship between older interviewers and attending community meetings. The empirical best linear unbiased predictors were also calculated for these models. There are only two outliers interviewer-question combinations are identified, one in the stone terrace question and one related to drought. The quantile graphs for these models are also shown in Figure 1 in the appendix.

With regard to the interviewer’s prior beliefs, the interviewer’s responses to the five subjective questions are included separately for each of the four objective questions. The sign and significance level of the impact of these effects are shown in Table 7. Of the twenty possible subjective-objective combinations, there are statistically significant results in two of the cases and two of the cases are not estimable as the models do not converge. As it seems unlikely that an interviewer’s prior beliefs about dispute resolution would impact the likelihood that the respondent indicates their plot has stone terraces, the significant relationships are most likely either spurious relationships or the interviewer’s beliefs are correlated with other characteristics that impact the quality of the enumeration.

4.4 Interaction Terms

Traditional relationships in Timorese society between age, gender, and deference create many potentially relevant combinations of characteristics. At the extreme, theory would support significant interaction term between a younger female respondent and an older male interviewer in which she is more likely to defer to his opinion. It is not possible to include this level of interaction terms because of dangers in over-fitting the model and having insufficient degrees of freedom for the estimation. Since there were more significant results with the first order effects related to gender, the interaction analysis focuses on those relationships, and on the interviewer’s opinion, as it is less well-studied than other relationships. Table 8 shows a sample of the possible interactions and the contrasts of the marginal predictions. The lines which note an individual covariate are the joint significance tests of the main effects, and the lines in which more than one covariate are noted are the joint tests of the two way interactions. In most cases, while the significance of the main effect remains robust to the inclusion of the interactions terms, at least in cases in which the effect was initially significant, the interaction terms are not significant.

An exception, however, is the interaction term between a female respondent and the interviewer's opinion. In three of the five models, Corruption, Law & Order, and Women's Rights, the interaction is strongly significant. These results are robust to the inclusion of the gender of interviewer and associated interactions. In two of the models, Corruption and Dispute Resolution, the interaction between a female interviewer, a female respondent, and the interviewer's opinion are also significant. The only question which shows no evidence of this is Land Titling. In no cases, however, is the interaction between a female interviewer and the interviewer's opinions significant. The conclusion that can be drawn is that female respondents to this survey were more likely to be influenced by the opinions of the interviewer than male respondents. This is true even if the interviewer is a woman. Also, these findings are robust to the inclusion of an education variable, and therefore the results are not driven in differences in education between genders. Overall these results suggest that female respondents are more likely to be influenced by the interviewers' opinions.

4.5 Salience

Salience may impact the magnitude of the interviewer effects if respondents are less susceptible to pressure if they believe strongly about a subject, or if interviewer are more likely to exert pressure if they feel strongly. To test this, I repeat the analysis for Women's Rights, separating the sample into male-only and female-only samples, and into male interviewer and female interviewer sample, under the hypothesis that women's rights issues would be less salient to men than women. The separate male and female respondent subgroups have roughly equal sample sizes ($n=638$ for men, $n=621$ for women). Using the fixed effects model, I find that the interviewer effects are jointly significant for the male-only sample ($p<0.0001$) but only weakly significant for the female-only sample ($p<0.0670$). Comparing the percentage of variance explained, again using a Nagelkerke test the overall interviewer effects explain a slightly higher proportion of the variance for the female-only sample, but overall the explanatory power of this model is much higher for the male sample ($R^2=0.250$) compared with the female sample ($R^2=0.162$). Therefore even the limited explanatory power of the interviewer effects contributes a large portion of the explanation for the female sample. Repeating the analysis with male and female interviewer subsamples, a Nagelkerke test shows that interviewer effects explain a higher portion of the variance in the female interviewer sample, 37.8 percent compared to 30.7 percent, but again the overall variance explained was lower (see Table 5). If these results are explained as Women's Rights being more relevant to urban education women than to the average female respondent, the results support the finding in Hermann (1983) that issues that are salient to interviewers are more likely to show interviewer effects while those important to respondent are less likely, but the results are again not conclusive.

5. Conclusions

Overall, the analysis finds strong evidence that interviewer effects are present in the data from the TLSLS-2X survey, but that these effects are stronger for subjective and potentially sensitive

questions compared to more objective questions. The respondent's characteristics explain the majority of the variation, even taking into account the fact that the explained variance of the interviewer effects is likely to be overstated due to confounding with geographic effects. The findings on the first order effects of the interviewer support the hypothesis that respondents are more likely to give pro-state answers to historically disenfranchised groups, but the evidence is not conclusive. Similarly the findings suggest women respondents are more susceptible to the influence of the interviewer's opinions on subjective questions.

The above results must be interpreted with the caveat that the study was one of convenience and is vulnerable to multicollinearity and omitted variable bias. The likely correlation between the interviewer's opinions and other covariates such as the interviewer's age, gender, education, and ethnic group raises the potential impact of multicollinearity of the estimates and significance generated with age, gender, and interviewer opinion, as well as omitted variable bias for education and ethnic group.

The results from this study, while not definitive, may provide a basis for future work more purposely designed to study interviewer effects. A design may employ an interpenetrating design, such as the one used in Schnell & Kreuter (2005), or a cross-classified estimation model, to separate the geographic and interviewer effects. The topics covered in a methodological study could also be selected to have more straightforward expectations and to have a sufficiently large sample size to explore a wider range of interactions. Even if future work is not forthcoming, the results highlight the need to consider more fully the impact of traditional cultural norms when conducting quantitative surveys in the developing world on topics that are outside the objective survey questions.

References

- Ballou, Janice. (1990) Respondent / Interviewer Gender Interaction Effects in Telephone Surveys. Presented at 1990 International Conference on Measurement Errors in Surveys. Tucson, AZ. November 11-14.
- Benstead, Lindsay J. (2010) Effects of interviewer gender and religious dress on survey responses: Findings from a Nationally-Representative Field Experiment in Morocco. In Annual Meeting of the Society of Political Methodology.
- Billiet, Jacques, and Geert Loosveldt. (1988) Improvement of the Quality of Responses to Factual Survey Questions by Interviewer Training. *Public Opinion Quarterly* 52 (2): 190–211.
- Blaydes, Lisa, & Gillum, Rachel. M. (2013) Religiosity-of-Interviewer Effects: Assessing the Impact of Veiled Enumerators on Survey Response in Egypt. *Politics and Religion*, 1-24.
- Bollinger, Christopher R., and Martin H. David. (2005) I Didn't Tell, and I Won't Tell: Dynamic Response Error in the SIPP. *Journal of Applied Econometrics* 20 (4): 563–69.
- Cannell, Charles F., Peter V. Miller, and Lois Oksenberg. (1981) Research on Interviewing Techniques. *Sociological Methodology* 12 (January): 389–437. doi:10.2307/270748.
- Catania, Joseph A., Diane Binson, Jesse Canchola, Lance M. Pollack, Walter Hauck & Thomas J. Coates. (1996) Effects of Interviewer Gender, Interviewer Choice and Item Wording on Responses to Questions Concerning Sexual Behavior. *Public Opinion Quarterly* 60: 345-375.
- Collins, Martin, and Bob Butcher. (1983) Interviewer and Clustering Effects in an Attitude Survey. *Journal of the Market Research Society*. <http://psycnet.apa.org/psycinfo/1983-22287-001>.
- Cotter, Patrick. R., Jeffery Cohen, & Philip B. Coulter. (1982) Race-of-interviewer effects in telephone interviews. *Public opinion quarterly*, 46(2): 278-284.
- Crowne, Douglas P., and David Marlowe. (1960) A New Scale of Social Desirability Independent of Psychopathology. *Journal of Consulting Psychology* 24 (4): 349.
- Davis, Darren W. (1997) The direction of race of interviewer effects among African-Americans: Donning the black mask. *American Journal of Political Science*, 309-322.
- Davis Rachel E., Mick P. Couper, Nancy K. Janz, Cleopatra H. Caldwell, & Ken Resnicow. (2010) "Interviewer effects in public health surveys." *Health Education Research* 25(1): 14-26.
- Duan, Naihua, Margarita Alegria, Glorisa Canino, Thomas G McGuire, and David Takeuchi. (2007) Survey Conditioning in Self-Reported Mental Health Service Use: Randomized Comparison of Alternative Instrument Formats. *Health Services Research* 42 (2): 890–907. doi:10.1111/j.1475-6773.2006.00618.x.
- Ellison, Christopher G., Michael J. McFarland, and Neal Krause. (2011) Measuring Religiousness Among Older African Americans: Exploring Race-of-Interviewer Effects. *Review of Religious Research* 53 (1): 65–84. doi:10.1007/s13644-011-0002-9.
- Flores-Macias, Francisco & Chappell Lawson. (2008) Effects of Interviewer Gender on Survey Responses: Findings from a Household Survey in Mexico. *International Journal of Public Opinion Research* 20 (1): 100-110.

- Finkel, Steven E., Thomas M. Guterbock, and Marian J. Borg. (1991) Race-of-Interviewer Effects in a Preelection Poll: Virginia 1989. *The Public Opinion Quarterly* 55 (3): 313–30.
- Fowler, Floyd J., and Thomas W. Mangione. (1990) *Standardized Survey Interviewing: Minimizing Interviewer-Related* Vol. 18. Sage.
- Freeman, J., and E. W. Butler. (1976) Some Sources of Interviewer Variance in Surveys. *Public Opinion Quarterly* 40 (1): 79–91. doi:10.1086/268269.
- Galla, John P., John D. Frisone, Linda R. Jeffrey & Eleanor P. Gaer. (1981) Effect of Experimenter's Gender on Responses to a Sex-Role Attitude Questionnaire. *Psychological Reports* 49: 935-940.
- Groves, Robert M. (1989) *Survey Error and Survey Cost*. New York: John Wiley.
- Hatchett, Shirley, and Howard Schuman. (1975) White Respondents and Race-of-Interviewer Effects. *Public Opinion Quarterly* 39 (4): 523–28. doi:10.1086/268249.
- Hermann, Dieter. (1983) Die Priorität von Einstellungen Und Verzerrungen Im Interview. Eine Methodenuntersuchung Anhand Der Daten Der Allgemeinen Bevölkerungsumfrage 1980. *Zeitschrift Für Soziologie* 12 (3): 242–52.
- Huddy, Leoine, Joshua Billig, John Bracciodieta, Lois Hoeffler, Patrick J. Moynihan, & Patricia Pugliani. (1997) The Effect of Interviewer Gender on the Survey Response. *Political Behavior*. 19(3): 197-220
- Johnson, Timothy P. & Robert W. Moore. (1993) Gender Interactions between Interviewers and Survey Respondents: Issues of Pornography and Community Standards. *Sex Roles* 28: 243-261.
- Johnson, Weldon T. & John D. DeLatmater. (1976) Response Effects in Sex Surveys. *Public Opinion Quarterly* 40: 165-181.
- Kahn, Robert L., and Charles F. Cannell. (1957) *The Dynamics of Interviewing; Theory, Technique, and Cases*.
- Kane, Emily W. & Laura J. Macaulay. (1993) Interviewer Gender and Gender Attitudes. *Public Opinion Quarterly Journal* 57: 1-28.
- Kennickell, Arthur B. (2004) Action at a Distance: Interviewer Effort and Nonresponse in the SCF. <http://www.federalreserve.gov/Pubs/OSS/oss2/papers/asa2004.5.pdf>.
- Krysan, Maria, and Mick P. Couper. (2003) Race in the Live and the Virtual Interview: Racial Deference, Social Desirability, and Activation Effects in Attitude Surveys. *Social Psychology Quarterly* 66 (4): 364–83. doi:10.2307/1519835.
- Landis, Judson R., Daryl Sullivan, & Joseph Sheley. (1973) Feminist Attitudes as Related to Sex of the Interviewer. *Pacific Sociological Review* 16: 305-314.
- Liu, M., and K. Stainback. (2013) Interviewer Gender Effects on Survey Responses to Marriage-Related Questions. *Public Opinion Quarterly* 77 (2): 606–18. doi:10.1093/poq/nft019.
- Lueptow, Lloyd B., Susan L. Moser & Brian F. Pendleton. (1990) Gender and Response Effects in Telephone Surveys about Gender Characteristics. *Sex Roles* 22: 29-42.

- Meng-Li, Yang & Ruoh-Rong Yu. (2008) The interviewer effect when there is an education gap with the respondent: Evidence from a survey on biotechnology in Taiwan. *Social Science Research* 37: 1321-1331.
- Olson, Kristen, and Andy Peytchev. (2007) Effect of Interviewer Experience on Interview Pace and Interviewer Attitudes. *Public Opinion Quarterly* 71 (2): 273–86.
- Phung, T. D., Hardeweg, B., Praneetvatakul, S., & Waibel, H. (2013) “Non-Sampling Error and Data Quality: What Can We Learn from Surveys to Collect Data for Vulnerability Measurements?” *World Development*, <http://dx.doi.org/10.1016/j.worlddev.2013.11.008>
- Rhodes, P. J. (1994) Race-of-Interviewer Effects: A Brief Comment. *Sociology* 28 (2): 547–58. doi:10.1177/0038038594028002011.
- Rice, S. A. (1929) Contagious Bias in the Interview: A Methodological Note. *American Journal of Sociology* 35 (3): 420–23.
- Rosen, Jeffrey, Joe Murphy, Andy Peytchev, Sarah Riley, and Mark Lindblad. (2010) The Effects of Differential Interviewer Incentives on a Field Data Collection Effort. *Field Methods*. <http://fm.sagepub.com/content/early/2010/09/01/1525822X10383390.abstract>.
- Samples, Tara C., Amanda Woods, Telsie A. Davis, Miesha Rhodes, Amit Shahane, and Nadine J. Kaslow. (2014) “Race of Interviewer Effect on Disclosures of Suicidal Low-Income African American Women.” *Journal of Black Psychology* 40(1): 27-46.
- Schaeffer, Nora Cate. (1980) Evaluating race-of-interviewer effects in a national survey. *Sociological Methods & Research*, 8(4): 400-419.
- Schnell, Rainer, and Frauke Kreuter. (2005) Separating Interviewer and Sampling-Point Effects. *Journal of Official Statistics* 21 (3): 389–410.
- Singer, Eleanor, Martin R. Frankel, and Marc B. Glassman. (1983) The Effect of Interviewer Characteristics and Expectations on Response. *Public Opinion Quarterly* 47 (1): 68–83.
- Suchman, Lucy, and Brigitte Jordan. (1990) Interactional Troubles in Face-to-Face Survey Interviews. *Journal of the American Statistical Association* 85 (409): 232–41.
- Tu, Su-Hao. & Pei-Shan Liao. (2007) Social Distance, Respondent Cooperation and Item Non-Response in Sex Surveys. *Quantity and Quality* 41: 177-199.
- Wallace, Roberty B., Frank J. Kohout, and Patricia L. Colsher. (1992) Observations on Interview Surveys of the Oldest Old. *The Oldest Old*, 123–34.
- West, Brady T., and Kristen Olson. (2010) How Much of Interviewer Variance Is Really Nonresponse Error Variance? *Public Opinion Quarterly* 74 (5): 1004–26.
- Wilk, Martin B. & Ram Gnanadesikan. (1968) Probability plotting methods for the analysis of data. *Biometrika* 55:1-17.

Appendix

Subjective Survey Questions

OPINION AND PERCEPTIONS OF THE LAW, RESPONDENT: HOUSEHOLD HEAD OR SPOUSE

Enumerator: If respondents says 'A', fill in 'A', if the respondent says 'B', fill in 'B'. These are the only two responses that should be offered to the respondent. If the respondent answers 'neither', fill in 'C', if the respondent answers 'both', fill in 'D'. Also note if the respondent doesn't know, or refuses to answer.

OPINIONS ABOUT JUSTICE

For each of the following pairs of statements, tell me which is closest to your view. There are no right or wrong answers to these questions, they are only about your opinions.

Question 1 [Corruption]

- A. Public officials are not paid enough, so it is acceptable for them to ask for additional payments beyond what is legally allowed in exchange for good service.
- B. Public officials serve the public and it is not acceptable for them to ask for additional payments beyond allowable rates.

Question 2 [Law & Order]

- A. The main responsibility for maintaining law and order in the community is with the police.
- B. The main responsibility for maintaining law and order in the community is with the *chefe de suco* and the community itself.

Question 3 [Women's Rights]

- A. A woman should be able to speak for herself in the traditional *adat* process.
- B. The head of the family should represent the household in the traditional *adat* process.

Question 4 [Dispute Resolution]

- A. It is better for non-violent disputes to be resolved within the community.
- B. The formal court system is designed to be fair to all citizens, and is the best place to take non-violent disputes.

Question 5 [Land Titling]

- A. A process of land titling would help to draw definite boundaries, and reduce conflict in the community.
- B. A land titling process would cause more problems than it solves by stirring up old problems and offering the opportunity to steal land with fake claims.

Objective Survey Questions

Community Meeting

I am now going to ask you about community decision making village/neighborhood. These meetings could be with the other organizations about projects and activities in the village / or with members of the community about village / or administration. Have you attended a decision making meeting in the past 12 months?

Siam Weed

Has Siam weed been a problem on this [PLOT]? Show card. (Responses were aggregated from the plot to household level.)

Stone Terraces

Does this [PLOT] have stone walls or stone terraces? (Responses were aggregated from the plot to household level.)

Drought

Since January of 2006, has your household experienced drought?

Table 1

Respondent and Interviewer Mean Characteristics (Weighted)

	<i>Respondents</i>		<i>Interviewers</i>	
	Mean	Std. Dev.	Mean	Std. Dev.
Female	0.49	0.50	0.39	0.50
Age	43.80	13.15	33.06	5.18
Ever Attended School	0.44	0.50	1.00	0.00
Completed Secondary	0.12	0.32	1.00	0.00
Years of Education	3.27	4.40		
Urban	0.23	0.42		
Log Per Capita Household Expenditure	3.38	0.55		

Table 2

Percentage of Respondents / Interviewers By Given Response (Un-weighted Sample Proportions)

	<i>Respondents (N=1396)</i>				<i>Interviewers (N=18)</i>			
	A	B	C*	D*	A	B	C*	D*
Corruption	40.5	56.0	1.8	1.7	25.0	71.4	0.0	3.6
Law & Order	58.6	32.4	0.8	8.2	75.0	25.0	0.0	0.0
Women's Rights	30.6	60.7	2.1	6.7	28.6	71.4	0.0	0.0
Dispute Resolution	46.1	41.8	4.2	8.0	21.4	71.4	0.0	7.1
Land Titling	70.9	20.0	2.5	6.5	71.4	25.0	0.0	3.6

Note: In addition to the options of A or B, respondents were technically allowed to select 'both,' option C, or 'neither,' option D, though these choices were not read to the respondent and enumerators were encouraged to probe for an A or B response.

Table 3

Distribution and Non-response for Objective Questions (Weighted)

	Yes	No	Missing
Community Meeting	49.2	50.4	0.4
Siam weed*	75.5	23.5	1.0
Stone terraces*	29.3	70.7	0.0
Drought	61.2	37.4	1.4

Note: Questions on Siam weed and stone terraces were asked only for those households with agricultural land.

Table 4 *Interviewer Fixed Effects*

	subjective					objective			
	corruption	law & order	women's rights	dispute resolution	land titling	community meeting	siam weed	stone terraces	drought
	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se
<i>Reference: 34 year old male</i>									
Male, 28 years old	-0.493 (0.442)	-1.918** (0.976)	-2.115** (1.078)	-2.725*** (0.841)	-1.309 (0.892)	-0.256 (0.462)	0.466 (0.774)	1.377 (1.168)	-1.415* (0.832)
Male, 28 years old	1.609* (0.856)	-1.834* (0.964)	2.989** (1.166)	0.463 (0.868)	0.805 (1.085)	1.310 (0.837)	-2.536** (1.007)	3.108** (1.494)	-1.809* (0.930)
Male, 29 years old	3.121*** (0.981)	-0.549 (0.834)	2.310** (1.160)	-0.304 (0.883)	-0.790 (0.942)	0.020 (0.733)	-0.152 (1.415)	3.448** (1.524)	-1.679** (0.771)
Male, 30 years old	1.520** (0.649)	0.182 (0.935)	0.304 (0.667)	0.040 (0.715)	0.502 (1.033)	-0.263 (0.476)	0.971 (0.809)	1.838* (1.100)	-0.530 (1.030)
Male, 35 years old	2.116*** (0.560)	-0.457 (0.826)	0.653 (0.652)	-1.224** (0.602)	-0.761 (0.915)	0.412 (0.681)	0.319 (0.789)	2.686** (1.201)	0.614 (0.986)
Male, 35 years old	1.993** (0.837)	-1.917** (0.908)	2.737** (1.132)	0.392 (0.852)	-0.065 (0.955)	1.148 (0.800)	-2.711*** (1.040)	2.681* (1.470)	-1.413 (0.934)
Male, 35 years old	-0.534 (0.925)	-1.662* (0.925)	1.966 (1.204)	0.173 (0.792)	-1.550* (0.942)	0.505 (0.710)	0.866 (0.911)	3.016** (1.498)	0.336 (1.029)
Male, 36 years old	-0.618 (1.207)	0.260 (1.554)	3.220*** (1.200)	-0.498 (1.621)	3.261** (1.522)	0.706 (0.564)	-0.178 (0.471)	-1.773 (1.659)	-0.855 (0.741)
Male, 39 years old	1.880* (1.071)	-1.740** (0.804)	3.510*** (1.215)	4.898*** (1.326)	0.216 (1.084)	0.309 (0.750)	-2.516** (1.178)	1.702 (1.595)	-1.596** (0.761)
Male, 40 years old	1.355 (0.913)	-2.453*** (0.937)	1.965* (1.186)	0.046 (0.779)	-1.894* (0.982)	1.035 (0.655)	-3.328*** (1.135)	3.032** (1.531)	-0.084 (1.017)
Female, 26 years old	2.419*** (0.626)	-0.965 (0.806)	3.465*** (0.830)	0.894 (0.824)	-0.488 (0.799)	0.458 (0.554)	-2.526** (1.020)	0.902 (1.464)	-2.545** (1.187)
Female, 26 years old	2.334*** (0.902)	-0.469 (1.221)	2.933*** (1.122)	0.759 (1.373)	0.433 (1.362)	-0.366 (0.995)	-1.942 (1.416)	1.591 (1.889)	-2.549* (1.447)
Female, 27 years old	-1.306 (0.988)	-0.256 (0.963)	1.513 (1.225)	0.066 (0.885)	1.132 (1.216)	1.251 (0.792)	-0.738 (1.204)	2.919* (1.533)	3.074** (1.408)
Female, 31 years old	2.465** (1.016)	-1.713** (0.826)	2.094* (1.190)	-0.41 (0.921)	-1.571* (0.930)	0.57 (0.743)	-3.362** (1.332)		-2.012*** (0.752)
Female, 31 years old	1.602* (0.846)	-0.931 (0.967)	3.446*** (0.845)	1.447 (0.984)	0.219 (0.919)	0.659 (0.751)	-4.347*** (1.293)	2.282 (1.572)	-2.311 (1.406)
Female, 37 years old	1.437 (0.894)	-0.631 (0.995)	3.567*** (1.219)	0.331 (0.906)	-0.311 (1.015)	0.427 (0.850)	-0.775 (1.050)	2.958** (1.485)	-2.625*** (0.917)
Female, 38 years old	2.023*** (0.632)	-0.117 (0.822)	0.637 (0.702)	-1.021 (0.667)	-0.129 (0.795)	-0.100 (0.503)	0.621 (0.971)	3.219*** (1.147)	-0.377 (1.075)
Female, 38 years old	3.209*** (0.872)	-1.087 (0.788)	1.737 (1.136)	0.449 (0.804)	-0.744 (0.930)	0.642 (0.702)	-1.818** (0.857)	3.001** (1.412)	-2.140*** (0.766)
Joint Significance F-stat	8.47	6.15	2.56	1.97	3.25	2.43	4.08	4.64	2.67
Prob > F	0.0000	0.0000	0.0023	0.0216	0.0001	0.0037	0.0000	0.0000	0.0016
Number of observations	1,213	1,133	1,135	1,095	1,130	1,249	1,087	1,096	1,177

Note: *** p<0.01, ** p<0.05, * p<0.1

Table 5

Variation explained by Interviewer Fixed Effects

	Adjusted R2 with interviewer effects	Adjusted R2 without interviewer effects	Percent of total variation from interviewer effects
<hr/>			
subjective			
corruption	0.160	0.093	41.9%
law & order	0.183	0.108	41.0%
women's rights	0.122	0.068	44.3%
dispute resolution	0.141	0.076	46.1%
land titling	0.174	0.101	42.0%
<hr/>			
objective			
community meeting	0.144	0.127	11.8%
siam weed	0.315	0.234	25.7%
stone terraces	0.273	0.216	20.9%
drought	0.255	0.224	12.2%
<hr/>			
women's rights			
men only sample	0.259	0.154	40.5%
female only sample	0.162	0.090	44.5%
male interviewers	0.199	0.138	30.7%
female interviewers	0.143	0.089	37.8%
<hr/>			

Table 6 *Mixed Effects Logit Results*

	subjective						objective		
	corruption	law & order	women's rights	dispute resolution	land titling	community meeting	siam weed	stone terraces	drought
	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se	coef/se
Respondent Characteristics									
female respondent	0.314** (0.135)	-0.025 (0.145)	0.360** (0.142)	-0.121 (0.139)	0.185 (0.161)	0.135 (0.128)	-0.070 (0.163)	-0.232 (0.161)	-0.246* (0.144)
<i>age (reference : age 30 to 39 years)</i>									
age below 30 years	-0.277 (0.203)	-0.060 (0.214)	-0.255 (0.217)	0.104 (0.214)	0.217 (0.254)	-0.566*** (0.197)	0.457* (0.251)	0.028 (0.258)	-0.065 (0.216)
age between 40 and 49	-0.141 (0.173)	0.154 (0.188)	-0.042 (0.183)	-0.215 (0.178)	0.137 (0.212)	0.123 (0.168)	0.069 (0.210)	0.007 (0.210)	-0.281 (0.186)
age greater than 50 years	-0.160 (0.188)	0.210 (0.202)	-0.035 (0.199)	-0.060 (0.192)	0.013 (0.222)	-0.201 (0.179)	0.203 (0.228)	0.100 (0.219)	-0.177 (0.207)
<i>education level (reference : no education)</i>									
primary	-0.159 (0.175)	-0.324* (0.186)	-0.057 (0.187)	0.200 (0.180)	-0.040 (0.209)	0.423** (0.166)	-0.005 (0.206)	0.069 (0.205)	-0.276 (0.185)
pre-secondary	0.078 (0.243)	-0.121 (0.257)	-0.196 (0.264)	-0.254 (0.250)	-0.096 (0.288)	0.449* (0.235)	-0.116 (0.298)	-0.507 (0.319)	-0.533** (0.260)
secondary	0.142 (0.220)	0.119 (0.249)	0.396* (0.228)	-0.081 (0.229)	-0.028 (0.261)	0.769*** (0.215)	-0.230 (0.276)	-0.670** (0.309)	-0.469** (0.236)
post-secondary	0.132 (0.482)	-0.147 (0.542)	0.891* (0.491)	-0.083 (0.507)	-0.731 (0.516)	1.459*** (0.512)	-0.241 (0.743)	-1.044 (1.085)	-1.798*** (0.636)
log per capita expenditure	0.301** (0.134)	0.355** (0.145)	0.043 (0.141)	-0.108 (0.135)	-0.044 (0.156)	-0.096 (0.127)	-0.442*** (0.167)	-0.173 (0.165)	-0.099 (0.147)
urban	-0.499*** (0.155)	0.047 (0.170)	0.021 (0.161)	-0.165 (0.161)	-0.274 (0.186)	-0.440*** (0.144)	-0.339* (0.183)	-0.445** (0.181)	-0.856*** (0.158)
Interviewer Characteristics									
female interviewer	0.305 (0.345)	0.555** (0.273)	0.088 (0.319)	-0.224 (0.313)	-0.113 (0.238)	-0.103 (0.144)	-0.188 (0.354)	0.443 (0.290)	-0.324 (0.206)
<i>interviewer age (reference 30 to 39 years)</i>									
age below 30 years	-0.248 (0.393)	0.431 (0.283)	-0.307 (0.343)	-0.380 (0.344)	0.645** (0.280)	0.075 (0.148)	0.400 (0.355)	-0.243 (0.286)	-0.426** (0.185)
age between 40 and 49	-0.363 (0.825)	-1.049* (0.560)	-0.242 (0.751)	-0.622 (0.712)	-0.854 (0.570)	0.546* (0.287)	-0.731 (0.700)	0.534 (0.522)	0.070 (0.450)
interviewer responded A	0.671* (0.403)	0.613* (0.320)	0.001 (0.443)	0.099 (0.448)	0.204 (0.353)				
Constant	-1.924*** (0.629)	-2.218*** (0.671)	-1.473** (0.640)	0.372 (0.615)	0.907 (0.735)	-0.240 (0.519)	2.467*** (0.766)	-0.549 (0.781)	0.485 (0.607)
n	1,213	1,133	1,140	1,075	1,094	1,254	1,087	1,100	1,237

Note: *** p<0.01, ** p<0.05, * p<0.1

Analysis also includes 12 district level fixed effects. Results not shown.

Table 7
Sign and significance of interviewer's priors on objective questions

		Sign and significance on objective question			
		community meeting	Siam weed	stone terraces	drought
Interviewer said A to:	corruption	- **	+	-	-
	law & order	-	+	+	+
	women's rights	+	+	+	-
	dispute resolution	+	+	- **	-
	land titling	+	+	n/e	n/e

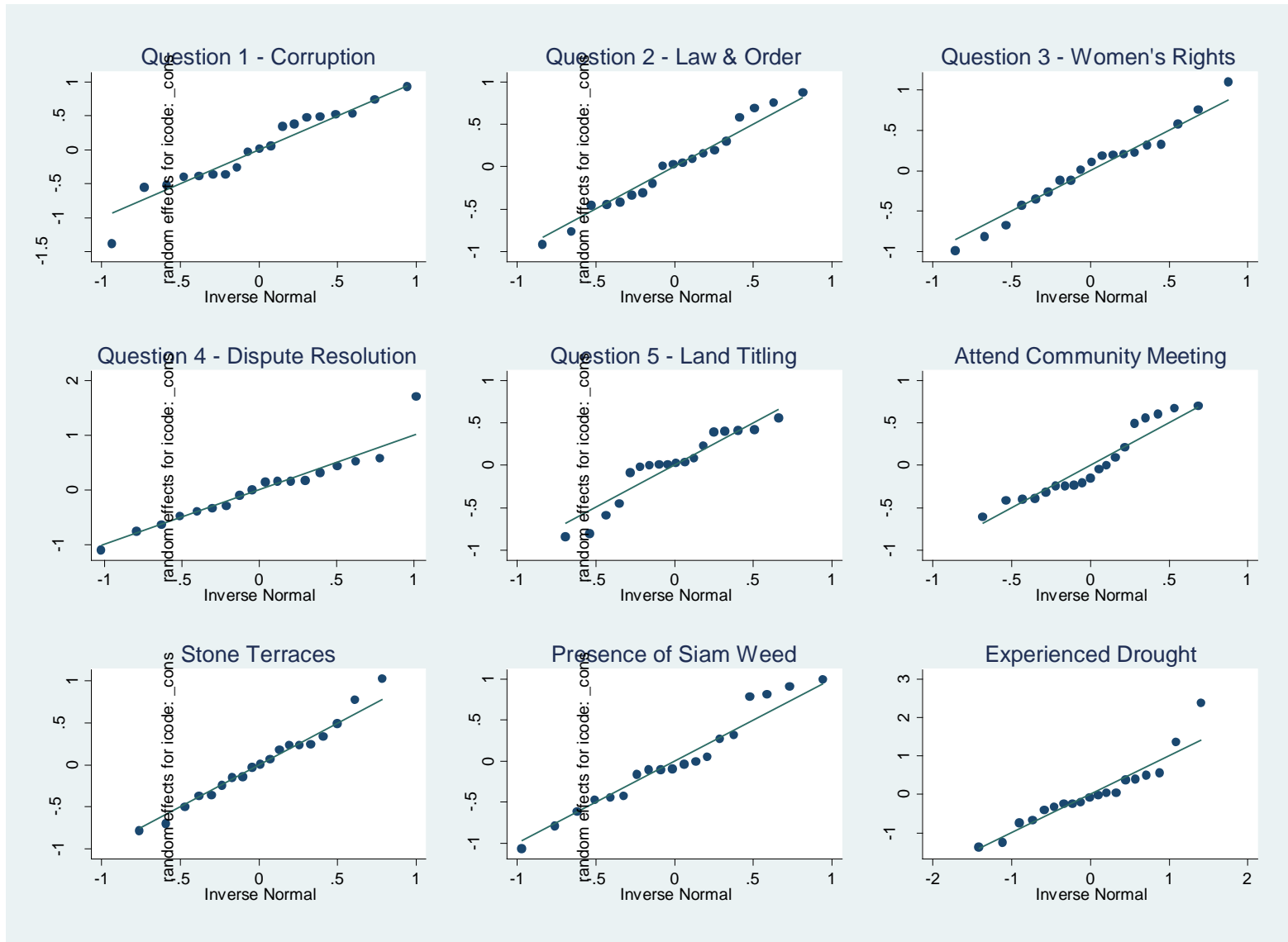


FIGURE 1 Empirical Best Linear Unbiased Predictors. Source: TLSLS-X (2009)

Table 8
Contrasts of marginal linear predictions

	df	corruption		law & order		women's rights		dispute resolution		land titling	
		chi2	P>chi2	chi2	P>chi2	chi2	P>chi2	chi2	P>chi2	chi2	P>chi2
Female respondent	1	5.27	0.022	0.01	0.927	6.25	0.012	0.84	0.361	1.04	0.309
Female interviewer	1	0.78	0.376	4.40	0.036	0.09	0.770	0.54	0.463	0.23	0.632
Female-Female interaction	1	0.01	0.944	2.62	0.105	1.55	0.213	0.27	0.604	0.56	0.454
Respondent Age Category	3	4.38	0.223	0.62	0.892	1.530	0.676	1.30	0.730	0.88	0.830
Interviewer Age Category	2	0.35	0.838	11.50	0.003	0.800	0.669	2.14	0.343	8.79	0.012
Age Category Interaction	6	9.09	0.168	8.44	0.208	3.660	0.723	6.50	0.369	4.51	0.609
Female interviewer	1	0.71	0.399	4.36	0.037	0.140	0.709	0.450	0.504	0.380	0.538
Respondent Age Category	3	2.02	0.569	1.05	0.790	1.420	0.701	2.350	0.503	0.970	0.809
Interaction	3	2.08	0.555	1.91	0.591	0.470	0.926	3.700	0.295	0.760	0.858
Female Interviewer	(not testable)										
Interviewer Opinion	1	3.13	0.077	3.00	0.084	0.010	0.936	0.010	0.914	0.280	0.597
Interaction	1	1.04	0.308	0.71	0.401	0.030	0.858	0.100	0.750	0.050	0.821
Female respondent	1	0.92	0.338	0.01	0.921	12.14	0.001	1.33	0.249	0.77	0.382
Interviewer Opinion	(not testable)										
Interaction	1	13.01	0.000	6.09	0.014	7.09	0.008	1.18	0.278	0.33	0.568
Female respondent	1	1.17	0.280	0.04	0.839	7.62	0.006	1.490	0.222	0.09	0.758
Female Interviewer	1	1.24	0.265	2.76	0.097	0.11	0.736	0.040	0.847	0.240	0.624
Female-Female interaction	1	0.71	0.400	0.97	0.325	0.16	0.687	4.360	0.037	1.81	0.179
Interviewer Opinion	1	3.46	0.063	3.45	0.063	0.01	0.923	0.050	0.822	0.220	0.641
Interaction (Female Respondent - Interviewer Opinion)	1	15.49	0.000	4.88	0.027	4.37	0.04	1.340	0.247	0.93	0.335
Interaction (Female Interviewer - Interviewer Opinion)	1	1.00	0.318	0.73	0.392	0.04	0.84	0.120	0.728	0.060	0.804
Interaction (Female Interviewer - Female Respondent - Interviewer Opinion)	1	8.51	0.004	0.21	0.646	0.41	0.52	10.750	0.001	1.54	0.215