Measurement When Respondents Lie

Daniel Rogger*

February 2015

Abstract

Survey methods typically elicit respondent characteristics directly from respondents. However, these responses may be intentionally biased or unintentionally contain measurement error. A response to both of these concerns is to invite assessments of respondent characteristics by co-respondents. It may be possible for a surveyor to identify a group of co-respondents whose average bias can be assumed to be lower than that of the respondent. Though co-respondents may individually make larger errors in their assessment of a respondent's characteristics, the ability to average those errors across multiple claims may lead to reductions in the overall bias. This paper develops these ideas and tests them using data on civil servants in Nigeria. It finds that measures of corruption based on co-respondent assessments perform far better than those based on respondent claims.

Keywords: Survey Methods, Corruption
JEL Classification: C83, D73

*Development Economics Research Group, World Bank. E-mail: drogger@worldbank.org. Web: www.danrogger.com.
I gratefully acknowledge financial support from the Economic and Social Research Council [ES/G017352/1], the Federal Government of Nigeria, IFS, the International Growth Centre [RA-2009-11-018], the Royal Economic Society and UCL. I thank the Presidency and the Office of the Head of the Civil Service of Nigeria for their support throughout data collection. I am grateful to Aidan Coville, Richard Murphy and Sami Stouli for useful discussions. All errors remain my own.

1 Introduction

Survey methods have almost universally relied on the direct elicitation of an individual's state. Individuals are asked by enumerators to characterise their state along multiple dimensions, from age, through income, to beliefs and experience of the world. However, this ignores the possibility that in households, communities, and organisations, there are other individuals surveyed, co-respondents, who may also have information on a respondents state which could either correct for any bias that the respondent introduces into their own response or increase the accuracy of measurement.

This paper looks at how greater use of co-respondents claims on respondents' states might improve the precision and efficiency of estimates of individual states of nature. This is particularly important when respondents may introduce large biases into their claims or make significant errors about their own characteristics. For example, when asked the extent to which a public official is involved in corrupt activities, the official may be averse to reporting the extent of the truth. Or when asked what their IQ is, individuals may make errors in their assessment simply because they cannot precisely measure it. Asking other individuals surveyed about a respondents characteristics may be a way to confront these issues. Other officials may be more willing to report the corruption of their colleagues than of themselves. Multiple assessments of a person's IQ from those who know them may be a better approximation *on average* than the single claim of the individual.[1]

Approaches to bias or measurement error in survey responses have typically targeted improvements in survey design, or statistical procedures to correct for the concerns. For example, recent responses to the bias induced by asking respondents about sensitive topics such as corruption or recreational drug use have focussed on the elicitation of direct responses through indirect questioning or the anonymisation of individual responses (see Krumpal et al (2014) for a review of the recent literature). One example of this is 'list randomization', wherein respondents are asked to count the number of statements in a list that are true for them. By giving a random subset of respondents a list with a key additional statement designed to capture sensitive behaviour, the surveyor can identify the proportion of the sample engaging in the sensitive behaviour without being able to identify the individuals engaged in it.

A distinct approach would be to use the knowledge of the other individuals in the survey (such as other members of the household, of the village, or of the organisation) to provide their own assessments of a respondents state. An advantage of using co-respondents is that there may be a group for whom the researcher is comfortable assuming the average bias is lower than that of the respondent. A disadvantage of using co-respondents is that they may have less accurate information over the true state of the respondent, and thus they may make larger unintended errors. However, as the number of co-respondents included in the joint assessment increases, for many error processes the mean of the average of these claims decreases, potentially to less than the expected error of the respondent themselves.

This approach would shift the focus of survey questions from an emphasis on the respondent to an emphasis on other respondents, or some aggregate of those others. For example, rather than asking public officials, 'On what proportion of the projects that you work on do you break the public service rules?', ask co-respondents, 'On what proportion of projects in your organisation do you observe (others/a specific group/a specific person) breaking the public service rules?' Rather than asking, 'How frequently do you use recreational drugs?', ask 'How frequently does (that individual/that group) use recreational drugs?'.

[1] This has been tested empirically by the psychology literature. There is a weak correlation between self-assessments of IQ and the IQ test results themselves (Paulhus, 1998), and co-respondents are more effective at assessing IQ, particularly teachers of pupils (Chamorro-Premuzic et al, 2009). Parents typically bias their child's IQ assessment upwards, though there are cultural determinants of that bias such as Japanese parents being too modest (Furnham and Fukumoto, 2008).

Though they have not been systematically analysed along the lines laid out in this paper, survey questions assessing others states have been fielded. In the field of educational psychology, psychologists have asked individuals to rank themselves within a group, requiring individuals to make assessments of others' states. That literature has found that accuracy is greatly improved when respondents have prolonged interactions (Kenny, 1991; Paulhus and Bruce, 1992), when the task is well defined and when the action has already occurred (Forest and Heath, 1996). In a distinct context, surveys of sanitation uptake have questions on the degree to which other individuals undertake sanitation activities.

Despite the existence of survey questions detailing the states of others in a small number of settings, in the main they have not been analysed in the setting laid out here. This paper contributes to the literature by highlighting the potential power of co-respondent assessments as a means to reduce bias and error in survey characterisations of respondents.

The rest of the paper is organized as follows. Section 2 lays out a theoretical justification for the use of co-respondent assessments in surveys. Section 3 then applies these ideas to data from the Nigerian civil service. Section 4 provides concluding comments and discussion.

2 Theory

2.1 Main idea

Let us suppose we are interested in an index θ that characterises some aspect of individuals in a population. A key aim in most respondent surveys is the recovery of θ_i, the state of nature for individual i. Whilst we might aggregate these quantities over multiple individuals, the heart of the survey exercise is frequently to identify the state of nature of each individual. Since this process is based on claims by respondents, the ambition is to minimise the distance between assessments of θ_i by respondents (or their co-respondents) and its true value.

Typically we aim to recover θ_i by using direct responses of individual i regarding her state of nature with respect to θ, $\hat{\theta}_{ii}$. Suppose that the response of individual i is made up of what individual i perceives to be the truth, $\widetilde{\theta}_i$, which is the truth plus a stochastic error term, ε_i, and a bias, $b_i \in (-\infty, \infty)$.[2] Whilst the deviation from θ_i due to the error is unintended by the respondent, the deviation from θ_i due to the bias is intended. The respondent's claim over θ_i is therefore,

$$\hat{\theta}_{ii} = b_p \left(\theta_i + \varepsilon_i\right) \tag{1}$$

where I have assumed that the proportion of the truth perceived by i that she reports is common across individuals and is equal to b_p.

Neither ε_i or b_p are observed by the econometrician. Thus, when estimating θ_i without knowledge of ε_i or b_p, our estimate is constrained to (2.1). When estimating features of θ across a group, such as the mean, direct responses give us,

$$\frac{1}{n}\sum_{i=1}^{n}\hat{\theta}_{ii} = \frac{b_p}{n}\sum_{i=1}^{n}\theta_i + \frac{b_p}{n}\sum_{i=1}^{n}\varepsilon_i \tag{2}$$

[2] I assume here the the individual does not try to correct for the error in their estimation or at least that we are dealing here with the error after any corrections.

where we defined the mean of θ as $\frac{\theta_1+\theta_2+\cdots+\theta_n}{n}$. At the limit, (2.1) becomes,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\hat{\theta}_{ii}=b_p\theta+b_p\lim_{n\to\infty}E\left[\varepsilon_i\right] \tag{3}$$

By limiting ourselves to recovering θ_i through the single direct response of the respondent, we limit ourselves to the bias and error of a single individual. If we expand the set of respondents that make claims on θ_i to co-respondents of i we may be able to find a group whose (potentially weighted) average estimate of θ_i may be closer to the true θ_i.

Assume that we have $n-1$ assessments of θ_i, with

$$\hat{\theta}_{ji}=b_{ji}\left(\theta_i+\varepsilon_{ji}\right) \tag{4}$$

where $\hat{\theta}_{ji}$ is the claim of co-respondent j regarding the value of θ_i, ε_{ji} is the error j makes over the true value of θ_i, and b_{ji} is the bias in the claim of j over i's state of nature. We can then take some statistic of these multiple claims on θ_i, such as the mean,

$$\frac{1}{n}\sum_{j\neq i}^{n-1}\hat{\theta}_{ji}=\frac{1}{n}\sum_{j\neq i}^{n-1}b_{ji}\theta_i+\frac{1}{n}\sum_{j\neq i}^{n-1}b_{ji}\varepsilon_{ji} \tag{5}$$

For each respondent i, we can construct a vector of co-respondent assessments of θ_i. In a survey setting, such a process may be too burdensome on respondents to be realistic. We can therefore use a variant of this principle by asking respondents to assess θ for all the other $n-1$ members of the group or organisation excluding them. This leads to a series of co-respondent claims of the following form,

$$\hat{\theta}_{ij}=\frac{b_o}{n-1}\sum_{j\neq i}^{n-1}\theta_j+\frac{b_o}{n-1}\sum_{j\neq i}^{n-1}\varepsilon_j \tag{6}$$

where we simplify by assuming that b_{ij} is same for all co-respondents, and equal to b_o, and that ε_{ij} has the same underlying distribution for all co-respondent assessments. In this case, we can estimate the group mean as follows,

$$\frac{1}{n}\sum_{i=1}^{n}\hat{\theta}_{ij}=\frac{b_o}{n\left(n-1\right)}\sum_{i=1}^{n}\sum_{j\neq i}^{n-1}\theta_j+\frac{b_o}{n\left(n-1\right)}\sum_{i=1}^{n}\sum_{j\neq i}^{n-1}\varepsilon_j \tag{7}$$

which in the limit becomes,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\hat{\theta}_{ij}=b_o\theta_j+b_o\lim_{n\to\infty}E\left[\varepsilon_j\right] \tag{8}$$

Note, even if there is no relative bias across respondents and co-respondents, but the respondent observes their state with noise, adding the responses of co-respondents increases the efficiency of the estimator of the respondents state.

To fix ideas, I will continue the discussion based around the construction of group or organisational means. These are relevant to any organisational characteristics, features of an individual in the group, or any team or group characteristic. Estimating across a vector of characteristics, rather than a single mean, follows the same logic.

By comparing across (2.1) and (2.1), we can see that there are two potential areas in which the use of co-respondent claims, rather than respondent claims, might reduce the distance between our assessment and the truth. First, ignoring the errors, when the average bias of co-respondents is lower than that of the respondent, $b_o < b_p$, the co-respondents claims will be better approximations of the true value of θ.

Second, any statistic we might construct that averages across multiple co-respondent claims will benefit from a potential reduction in the average error. Whilst the expected size of the relative errors will depend on the exact nature of the underlying error process, in a large class of plausible cases, averaging across stochastic draws from a distribution will drive the estimate towards a low relative mean.

A simple example illustrates how having multiple assessments of a respondent's characteristic may reduce the overall bias. Ignoring the bias, suppose that co-respondents were prone to making errors over the true state of nature for i that were on average twice as large as that of respondent i. Thus, $\sigma_o^2 = 2\sigma_p^2$, and there were no other differences between the error processes of respondents and co-respondents. Since $E\left[\frac{1}{n-1}\left(\theta_{1i} + \theta_{2i} + \cdots + \theta_{(n-1)i}\right)\right] = 0$ and $E\left[\left(\frac{1}{n-1}\left(\theta_{1i} + \theta_{2i} + \cdots + \theta_{(n-1)i}\right)\right)^2\right] = \frac{\sigma_o^2}{n-1}$, the co-respondents joint expected assessment of θ would be more accurate than the respondent's after the third co-respondent assessment was included in the analysis. Similarly, if co-respondents were liable to make q times the error of respondents, by the inclusion of the $(q+1)$th co-respondent assessment, their accuracy would be greater in expectation than that of a single respondent.

Each of these steps requires assumptions to be made on the underlying relative claims of respondents and co-respondents. The utilisation of co-respondent claims rests on the relative sum of bias and error being smaller for the group of co-respondents utilised for estimating the respondent's state. However, we can investigate the impact of bias and measurement error separately. The bias caps the coefficient on corruption to a finite level, whilst measurement in a classical errors-in-variables setting reduces the absolute value of the coefficient. This second effect is more extreme at lower numbers of observations, whilst the bias effect is constant.

To see this diagrammatically, Figure 1 describes the expected path of the coefficient on the measures defined by respondent and co-respondent analysis respectively as we increase the number of individual responses used to make up the two measures. The horizontal axis is the number of respondents or co-respondents utilised in the construction of the measure, and the vertical axis is the coefficient. We will investigate a setting in the empirical section where the true parameter is negative, and so the graph echoes this with a true θ less than zero. As the number of individuals used to construct the measure increases, the absolute size of the coefficient rises, as the relative measurement error decreases (the measure is being created using the information from a larger number of individuals, reducing average measurement error). We therefore see the two sets of coefficient move towards the true value of θ. However, the bias caps the coefficient on the respondent-based measure at an artificially low level, with the co-respondent coefficient converging to the truth.

In this example, the co-respondent measure dominates the respondent measure at all levels, but it is feasible that in some settings the co-respondent coefficient curve crosses the respondent coefficient curve at a point $n^\star$, where the number of observations utilised to create the measures is such that the mean relative deviation of the co-respondent measure becomes less than that of the respondent measure.

With sufficient sample size and appropriate assumptions on the error terms, comparison of the respondent and co-respondent means is a direct measure of the relative bias. In Figure 1, the vertical space labelled as (1) is a deviation in the coefficient from its true value due to measurement error, whilst the vertical space labelled (2) is a deviation arising from the relative biases, and assuming a zero-bias in the co-respondent measure, is a measure of the bias of respondents over their own states. We will be able to empirically replicate this logic in the empirical section.

2.2 Extensions

This paper assesses the tradeoff between the reduction in bias and the additional noise that arises from asking co-respondents about respondents states of nature. So far, the discussion has taken all of the co-respondents as equally valid contributors to the assessment of a respondents state. It may be that some co-respondents are more informed about a respondent than others, and efficiency would dictate that we weight their claims over the respondent's state more highly than those of others. So long as we can obtain measures or proxies of how closely co-respondents know respondents (specifically along the margin being assessed in the survey), we can weight different co-respondents claims differentially.

We can weight a co-respondents responses in two ways. First, we may believe that the individual co-respondent has superior information across all other respondents (they have a greater appreciation of the group common across respondents). We can see this as a common weighting, w_i. Second, we can vary the weighting by respondent-co-respondent pairs, where we think that some co-respondents have superior knowledge about some respondents but not others. We can see this as a vector of $i - j$ pairs, w_{ij}. This implies that (2.1) becomes,

$$\frac{1}{n}\sum_{i=1}^{n}\hat{\theta}_{ij} = \frac{b_o}{n\,(n-1)}\sum_{i=1}^{n}w_i\sum_{j\neq i}^{n-1}w_{ij}\theta_j + \frac{b_o}{n\,(n-1)}\sum_{i=1}^{n}w_i\sum_{j\neq i}^{n-1}w_{ij}\varepsilon_j \tag{9}$$

In the corruption setting, when a co-respondent is making claims at the organisational level ('my colleagues at this organization are corrupt on x% of projects'), we can weight their responses by a proxy of how well they know the organization. For corruption assessments of individuals by individuals, it could be 'what proportion of projects that your colleague works on do you work on with him/her?' Clearly there needs to be a balance between improved knowledge of another individual, and the degree to which proximity leads to an increase in bias due to self-incrimination or another reason.

So far we have looked at a case in which everyone is similarly biased about their own responses and not about others. Where there are reasons to suspect this may not be true, we can again vary the weightings over bias by co-respondent, b_i, and by respondent-co-respondent pairs, b_{ij}. This leads to an expression of the form,

$$\frac{1}{n}\sum_{i=1}^{n}\hat{\theta}_{ij} = \frac{1}{n\,(n-1)}\sum_{i=1}^{n}b_iw_i\sum_{j\neq i}^{n-1}b_{ij}w_{ij}\theta_j + \frac{b_o}{n\,(n-1)}\sum_{i=1}^{n}b_iw_i\sum_{j\neq i}^{n-1}b_{ij}w_{ij}\varepsilon_j \tag{10}$$

There may be concerns over the separability of bias and co-respondent knowledge. If a co-respondent is close enough to a respondent that she has superior information about her, she may also feel committed to limit truth-telling about her.

3 Empirical Evidence

To validate the efficacy of the method laid out above, we require a data set that has respondent assessments of their own state of nature, co-respondent assessments of others state of nature, and credible independent assessment of the true value of the state of nature. A motivating theme for the argument laid out here is that self-incrimination may be a source of serious bias in self-reporting. Thus, an optimal setting for an empirical investigation would be one in which respondents fear incriminating themselves.

Such a setting is the discussion of corruption in public agencies in the Federal Government of Nigeria. To investigate the issues laid out above empirically, I exploit data from a unique data set that surveys civil servants in Nigeria on both their and others corruption, and contains independent assessments of the infrastructure projects they work on; projects that suffer from any corruption they might be involved in.[3]

3.1 Data

The data I use studies 63 organizations in the Nigerian federal civil service, including government ministries and other federal agencies. I measure the claims civil servants in these 63 organizations make about the corruption they either come under pressure to undertake *themselves*, or that they observe *others* involved in.

I measure civil servants' claims over corruption through a survey fielded to a representative sample of 4100 civil servants, corresponding to 13% of all bureaucrats in the 63 organizations I study. Civil servants were asked to 'Think about recent projects and/or programmes you worked on for this organisation. In what proportion of the projects have you had to face the following difficult challenges?' They were then asked to consider this statement for a number of options, including 'I was put under pressure to divert some of the funds' and 'I observed others breaking service rules for their own benefit'. The first of these is a measure of the individual experience of corruption akin to (2.1), that could be perceived to be self-incriminating. The second of these is a measure of the perception of others corruption akin to (2.1).

At the organisational level, bureaucrats state that they are put under pressure to engage in corrupt activities on 17% of projects. However, they claim that they observe others engaging in corrupt activities on 37% of projects. This is remarkably close to the number of projects in our data that do not ever get started, as will be described below. Thus, the levels averages seem to imply that civil servants are more willing to report on the corruption of others than of themselves.

To verify the validity of these claims, I assess the extent to which the organizations implement 4700 small-scale infrastructure projects. Projects are of 11 types common to the Nigerian public sector, including construction (boreholes, buildings etc.) and non-construction types (procurement, training etc.). These public projects are managed by the civil servants I interviewed and though not the only source of corrupt rents, they are the main source of public funds managed by the civil servants. To assess the extent to which these projects were hindered by corrupt practices, I worked with a set of engineers fully independent of government to trace the implementation of each of these projects. Monitoring teams visited project sites 18 months after the infrastructure projects were approved and recorded whether the project had started, and its stage of completion.[4]

[3] Rosenfeld et al (2014) note how limited access to suitable sensitive information has limited the empirical study of asking sensitive questions, emphasising the contribution of this data set.

[4] Evaluations of the OPEN process indicate it successfully achieved its aims (Eboh 2010, Dijkstra et al 2011). To ensure the accuracy of monitoring reports, the Presidency put in place a system of checks and balances. First, a centralized team of technocrats monitored the evaluation teams, providing them with training and opportunities for standardization of their

My outcome variable is a continuous zero to one measure of project completion rates. Whilst all projects are intended to have been completed by the time the monitoring teams reached the projects, the average completion rate was just over 50%. The majority of the variation arises from some projects being starved of funding within the organisation. Though funding was budgeted for all the projects, 37% of projects are never started, whilst 31% of projects are completed fully.

For all of the organizations under study I collected data on their budgets, staffing and management practices. Given that corruption might vary across project and organisational observable characteristics, I will condition on these in the regressions, as well as on a series of noise controls that measure characteristics of the interviews, such as the day of the week and its duration.

3.2 Results: Main idea

To test the relative efficacy of respondent versus co-respondent survey techniques, we can benchmark the degree to which claims of organisational corruption match independent measures of that corruption. The measures of project completion are a major avenue for corruption in the Nigerian civil service. However, being one of a number of such avenues (others being rents from recurrent expenditures and extortion from citizens), we should see them as a lower bound on corruption. The coefficient on a successful measure of corruption should therefore approximate or exceed -1, implying that a percentage point increase in claims of corruption leads to a percentage point decrease in the level of public projects implemented.

My empirical specification has as its unit of observation project i of type j in organization n. I estimate the following OLS specification, where y_{ijn} is the project completion rate and $corrupt_n$ is the measure of corruption in the organization,

$$y_{ijn} = \beta_0 + \beta_1 corrupt_n + \beta_2 PC_{ijn} + \beta_2 OC_n + \lambda_j + \epsilon_{ijn} \tag{11}$$

$PC_{ijn}$and OC_n include project and organizational characteristics respectively. As many organizations implement project type j, I control for project type fixed effects λ_j. Standard errors are clustered by project type-organization. The coefficient of interest across specification is β_1 as we vary the measure of corruption from that built using direct responses to that built using co-respondent assessments. The core assumption underlying analysis is that the bias exhibited by public officials when describing the pressures they are under for corrupt acts is greater than the bias they exhibit when making claims over other public officials in the organisation. This is the core assumption of the analysis and it cannot be tested.

Table 1 presents our results and shows that the way corruption is measured matters. Column 1 presents the results of running specification (3.2) using a measure of corruption based on direct responses. The coefficient is negative, but well below -1, and insignificant at the normal levels. Taking this result at face value, a researcher would imply that corruption does not significantly impact the quality of an organisation's outputs.

Column 2 of Table 1 presents the results of specification (3.2) using a measure of corruption based on co-respondent assessments. Now the coefficient is (correctly) well above 1 in absolute terms, and the coefficient is statistically significant at the 1% level. Using this method of constructing a corruption measure,

methods at national conferences. Second, evaluators were asked to provide material, photographic, or video evidence to support their reports. Third, the national teams and Presidency performed random checks on evaluated sites, all of which were consistent with the findings of OPEN monitors.

Table 1: Respondent and Co-Respondent Measures of Corruption

	(1) Pressure	**(2) Observation**	**(3) Weighted**
Corruption measure	-0.70	-1.22***	-0.89***
	(0.43)	(0.33)	(0.32)
Controls		Project and Organizational	
Fixed Effects		Project Type	
Observations (clusters)		4721 (201)	

*** denotes significance at 1%, ** at 5%, and * at 10% level. Standard errors in parentheses are clustered by project type within organization. Project type fixed effects relate to whether the primary project classification is a financial, training, advocacy, procurement, research, electrification, borehole, dam, building, canal or road project. Project controls include the budget, whether it is new or rehabilitation project, and a measure of its technical complexity. Organizational controls include the logs of number of employees, total budget, and capital budget, the share of the workforce with degrees, with postgraduate qualifications, and management practice indices related to autonomy, incentives/monitoring and other practices. Noise controls include interviewer dummies, indicators of the seniority, gender, and tenure of the managers who responded, the day of the week and time the interview was conducted, a dummy for whether the interview was conducted during Ramadan, the interview duration, and an indicator of the reliability of the information as coded by interviewers.

a researcher would imply that corruption is an important determinant of organisational productivity in the Nigerian public sector.

The regression results are a first validation of the principles laid out in this paper. However, as described in section (2.1), we can use this data for other tests of validation. For the results above, I utilised all of those individuals surveyed to construct the measures of corruption. To test the impact of increased observations on our estimates, I can take a random sub-sample of the officials I interviewed from the total pool, and estimate the coefficient on these. I can do this for 1%, 2%, 3% and so on of my pool of officials and plot the dynamics of the coefficients as I increase the sample.

To avoid a lucky draw, I draw 100 random numbers to act as seeds to generate a further 100 random numbers each, and repeat the above process 100 times. I then average across these results, and plot the average coefficients in Figure 2. The coefficients relating to measures based on respondent assessments are those in blue, whilst those based on co-respondent assessments are those in red.

The results closely mirror what we would expect if the estimates were suppressed by both bias and measurement error. As I include a larger proportion of the officials interviewed in the construction of the corruption measures, the compression of the estimate is reduced. The measure of corruption based on direct responses converges to a lower value than that based on co-respondent assessments, reflecting the difference in bias.

3.3 Results: Extensions

In the OPEN context, we have claims on how corrupt colleagues at the organization are. In section 2.2, I argued that individuals who had better information over the organisation should be given a greater weighting in the analysis. Here, to weight the responses of co-respondents, we therefore require an indicator of how well the co-respondent can detect the corruption of others in the civil service. For this example, I use the number of years the individual respondent has been at within the civil service. This is akin to w_i in equation (2.2). I do not have data suitable to estimate variable weights at the respondent-co-respondent level, or to estimate variable biases.

The results are displayed in Column 3 of Table 1. We see that the coefficient on the co-respondent measure is still close to 1 and statistically significant at the 1% level, but is not as large as that without weighting. It may be that those who have been longest in the service know it best, but are also most reticent to discuss sensitive matters more generally.[5] This combination of bias and weighting was briefly discuss in the theory section, and is an area for further empirical exploration with data rich enough to separate weighting and bias.

4 Discussion

This paper has proposed a novel method for designing survey questions that relies on the information that co-respondents have about respondents. Rather than asking individuals directly to characterise themselves on sensitive topics, we ask others in the survey that know them to assess their true state of nature. This has the benefit of allowing us to find a group of co-respondents who may have a lower bias in responding, and by averaging across many co-respondents, possibly a lower average error as well.

The determinants of bias will vary across individuals, but the design of questions for co-respondents should take into account how closely a question framing is to the co-respondents own identity. A co-respondent may introduce little bias into a discuss of other members of their village, but may strongly identify with the *larger* conception of clan, and introduce a greater bias in response to questions related to clan. In the same way as questions have traditionally aimed to be framed in a way respondents find most palatable, questions designed for co-respondents, though providing a greater degree of freedom, may be subject to their own constraints.

Similarly, framing questions such that they facilitate co-respondents making the smallest errors possible would be useful in increasing the benefit of this method of investigation. Given the novelty of co-respondent assessment, there is considerable design features that require testing and further research.

References

CHAMORRO-PREMUZIC, TOMAS, ADRIANE ARTECHE, ADRIAN FURNHAM AND NADINE TRICKOT (2009) "Assessing pupils' intelligence through self, parental, and teacher estimates," Educational Psychology 29: 83-97

DENISSEN, J. J. A., SCHöNBRODT, F. D., VAN ZALK, M., MEEUS, W. H. J. AND VAN AKEN, M. A. G. (2011) "Antecedents and consequences of peer-rated intelligence." Eur. J. Pers., 25: 108–119

DIJKSTRA.G, B.AKANJI, C.HIDDINK, S.SANGARABALAN AND F.DE MEVIUS (2011) Mutual Interests - Mutual Benefits: Evaluation of the 2005 Debt Relief Agreement Between the Paris Club and Nigeria, The Hague: Ministry of Foreign Affairs of the Kingdom of the Netherlands.

EBOH.E (2010) MDGs-based Planning in Africa: Lesson, Experiences and Challenges: A Case Study of Nigeria, Addis Ababa: United Nations Economic Commission for Africa.

[5] The weighted version of the pressure measure of corruption implied a coefficient of -0.16, a standard error of 0.41 and a corresponding p-value of 0.693.

FOREST, J.J. and C. HEATH (1996) "The Evaluation Gap in Performance Perceptions: Illusory Perceptions of Group and Individuals," Journal of Applied Psychology 81(4): 369-379

FURNHAM, ADRIAN AND SARA FUKUMOTO (2008) "Japanese parents' estimates of their own and their children's multiple intelligences: Cultural modesty and moderate differentiation," Japanese Psychological Research 50: 63–76

KENNY, D.A. (1991) "A general model of consensus and accuracy in interpersonal perception," Pyschological Review 98: 155-163

KRUMPAL, IVAR, BEN JANN, KATRIN AUSPURG AND HAGEN VON HERMANNI (2014) "Asking Sensitive Questions: A Critical Account of the Randomized Response Technique and Related Methods" in UWE ENGEL, BEN JANN, PETER LYNN, ANNETTE SCHERPENZEEL, PATRICK STURGIS (2015) "Improving Survey Methods: Lessons from Recent Research" New York: Routledge.

PAULHUS, DELROY and M.N. BRUCE (1992) "The effect of acquaintanceship on the validity of personality impressions: A longitudinal survey," Personality Processes and Individual Differences 63: 816-824

PAULHUS, DELROY (1998), Self-Report Measures of Intelligence: Are They Useful as Proxy IQ Tests?. Journal of Personality, 66: 525–554

ROSENFELD, BRYN, IMAI KOSUKE AND JACOB SHAPIRO (2014) "An Empirical Validation Study of Popular Survey Methodologies for Sensitive Questions" Princeton University Working Paper.

Figure 1: Theoretical validation of co-respondent methodology

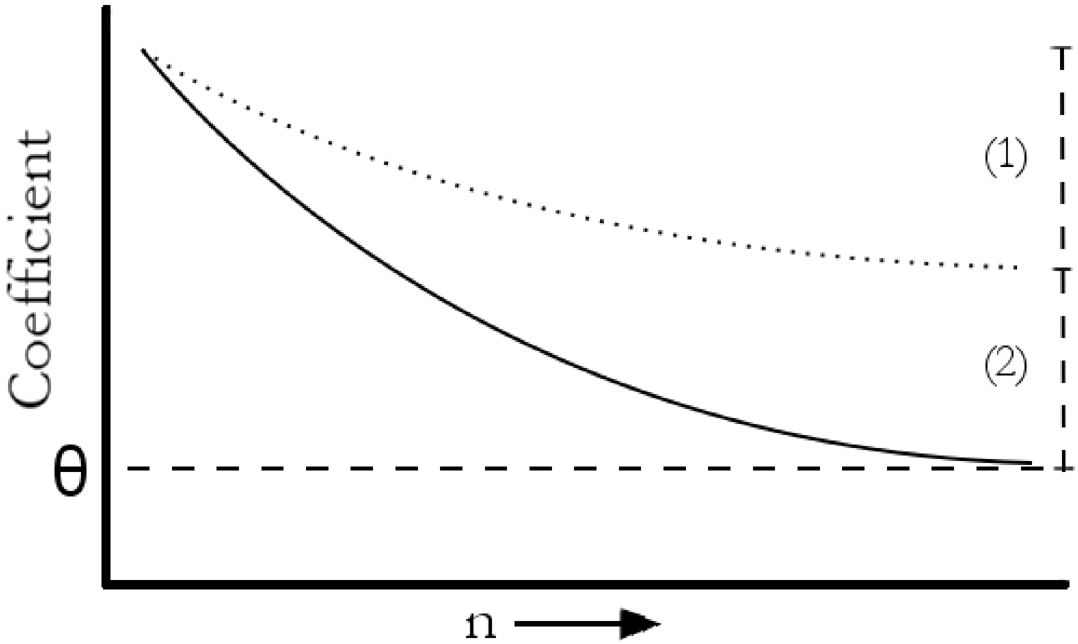

Figure 2: Empirical validation of co-respondent methodology