

Gauthier, Bernard; Lesné, Frédéric

Working Paper

Measuring corruption in presence of reticent respondents: Theory and Application

FERDI Working Paper, No. P207

Provided in Cooperation with:

Fondation pour les études et recherches sur le développement international (FERDI),
Clermont-Ferrand

Suggested Citation: Gauthier, Bernard; Lesné, Frédéric (2017) : Measuring corruption in presence of reticent respondents: Theory and Application, FERDI Working Paper, No. P207, Fondation pour les études et recherches sur le développement international (FERDI), Clermont-Ferrand

This Version is available at:

<https://hdl.handle.net/10419/269488>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Measuring corruption in presence of reticent respondents: Theory and Application

Bernard GAUTHIER, Frédéric LESNÉ.

➔ BERNARD GAUTHIER, HEC Montréal, Canada

➔ FRÉDÉRIC LESNÉ is a PhD candidate in Development Economics , CERDI - Université Clermont Auvergne. Corresponding author. **Email: frederic.lesne@etu.udamail.fr**

Abstract

We propose an alternative two-step method to correct the bias in corruption indicators associated with the presence of reticent respondents in firm surveys. Our method uses indirect and direct questions on bribery payments. An indirect question is used to identify reticent respondents, which makes it possible to adjust responses to a direct question assessing bribery activities. A simple theoretical model is presented to examine respondent behavior to sensitive questions using direct and indirect formulations. Applying the two-step method to a survey of 382 newly created firms in Madagascar, we find that the frequency of bribery measured by the standard way of measuring corruption activities ignoring reticence is underestimated by 47 percent.

Measuring corruption in presence of reticent respondents: Theory and Application

Bernard Gauthier*, Frédéric Lesné†

November 30, 2017

Abstract

We propose an alternative two-step method to correct the bias in corruption indicators associated with the presence of reticent respondents in firm surveys. Our method uses indirect and direct questions on bribery payments. An indirect question is used to identify reticent respondents, which makes it possible to adjust responses to a direct question assessing bribery activities. A simple theoretical model is presented to examine respondent behavior to sensitive questions using direct and indirect formulations. Applying the two-step method to a survey of 382 newly created firms in Madagascar, we find that the frequency of bribery measured by the standard way of measuring corruption activities ignoring reticence is underestimated by 47 percent.

1 Introduction

A well-known issue with firm-level surveys on bribery is that respondents are not always candid when answering sensitive questions about this illegal and socially undesirable behavior. Firm owners and managers may believe that questions about bribery are too intrusive, or that providing an honest answer may harm them. They may be reluctant to report the true extent of their involvement in corrupt acts because they are afraid of sanctions from authorities or because they fear social disapproval from the interviewer.

Following Azfar and Murrell (2009), a reticent respondent could be defined as “one who gives knowingly false answers with a nonzero probability when honest answers to a specific set of survey questions could lead to the inference that the respondent might have committed a sensitive act”, where the “sensitive act” is for instance the payment of a bribe to a public official.

*HEC Montréal, Canada

†Centre d’Études et de Recherches sur le Développement International, Université Clermont Auvergne, France. Corresponding author. Email: frederic.lesne@etu.udamail.fr

Among the determinants of reticence are respondents' characteristics, including whether they have anything to hide (Tourangeau and Yan, 2007 ; Schaeffer, 2000), as well as to the sociopolitical environment in which firms operate (Jensen *et al.*, 2010).

As reticent respondents' answering behavior to sensitive questions is systematically different from that of candid respondents, having the tendency to under-report socially undesirable behavior, bribery indicators are likely to be downward biased the more reticent surveyees in the firms sample.

To reduce reticence and potential biases in corruption indicators, researchers have developed various reticence-reducing techniques applied to the design and conduct of surveys on this sensitive topic. One of the most commonly used technique to reduce reticence is to ask sensitive questions indirectly (Fisher and Tellis, 1998). With this approach, survey respondents are not asked directly about their own experience regarding the behavior subject to under-reporting, but indirectly about their perception of the behavior of individuals or entities similar to them.

Indirect questioning gives respondents the opportunity to distance themselves from their answers, allowing firm executives to acknowledge the existence of bribery in their sector, without requiring them to admit to having personally paid any bribes.

Asking indirect questions about bribery is now standard practice in firm-level surveys. The World Bank's *Enterprise Surveys* for instance, one of the main sources of micro-level corruption data, make use of several indirect questions on bribery. One question in particular asks firm owners and managers to estimate the amount of bribes paid annually by firms like theirs to "get things done".

However, as emphasized by Clarke *et al.* (2015), respondents who are reticent to report how much bribes they pay may also be reticent to estimate what other firms pay in bribes. Azfar and Murrell (2009) found that respondents are equally reticent to answer direct and indirect questions about bribery, suggesting that "asking about others does not decrease reticence".

To reduce reticence to honestly answer sensitive questions, another proposed survey technique is the randomized response questioning (RRQ) method. First introduced by Warner (1965), the RRQ method adds a random component to survey answers, allowing respondents to answer honestly while keeping hidden their true behavior. In its most basic form, the method asks respondents to toss a coin each time before answering a series of sensitive questions in which a positive response implies that they have had an actual experience of the sensitive behavior subject to under-reporting. The respondent's answer hence partly depends on the outcome of the toss. If the coin comes up heads, respondents are asked to respond sincerely, while if tails come up, respondents must always answer positively, whatever their true answer is. Nobody except the respondent knows if a positive answer is the result of the coin tossing or an admission of the sensitive behavior by the

respondent. Given that the two sides of a fair coin have each a 50 percent probability of being drawn, it is however straightforward to compute the actual frequency of genuine positive answers in the sample of respondents by subtracting from the actual number of “Yes” answers the additional positive answers resulting from the probability of tossing the coin heads.

However, recent meta-analyses of RRQ studies have concluded that the RRQ technique has had limited success in reducing reticence in surveys (Lensvelt-Mulders *et al.*, 2005). Respondents appear to be generally reluctant to answer randomized response questions properly, with a non-trivial number of survey respondents not following the procedures and choosing to answer “No” even if their toss of the coin requests them to answer “Yes”.

Given these limitations, Azfar and Murrell (2009) proposed to use the RRQ procedure not to reduce reticence but to identify and eliminate reticent respondents from the estimation of bribery indicators. However, the resulting loss of information is an important limitation of this approach.

A variant of this approach was proposed by Kraay and Murrell (2016) who identified reticent respondents based on a statistical model of response behavior using a combination of RRQ and conventional survey questions. Applying the method to a firm-level survey in Peru, they estimated that the proportion of firms experiencing corruption is about twice as large as the estimate not accounting for reticence.

While the idea of using RRQ to identify reticent respondents is a clear breakthrough in survey-based corruption measurement, the method has several important drawbacks. First, the procedures requiring respondents to toss a coin before answering each RRQ, need to be plainly explained individually to respondents, increasing the length and complexity of the survey, introducing a higher risk that respondents decide not to complete the survey or choose to respond without having properly understood the rules. Clausen *et al.* (2010) reported that in the opinion of surveyors, about 14 percent of respondents to a firm-level survey in Nigeria did not properly understand randomized response procedures, which resulted in most of those respondents answering “No” to all randomized questions¹.

Furthermore, the RRQ approach applied to corruption usually ask respondents about sensitive behavior unrelated to corruption itself². Some

¹Clausen *et al.* (2010) discarded from their analyses respondents who were perceived by surveyors as not having understood the randomized response procedure. As a consequence, confusion over randomized response instructions on the part of survey respondents causes a non-negligible loss of information. It is actually possible that some respondents are reluctant to answer randomized questions because they do not understand the instructions or the objective of the procedure without being reluctant to honestly answer simpler questions asking them about their experience or perception of corruption.

²An example is Azfar and Murrell (2009) in which one of the seven randomized questions used to detect reticence to answer questions about corruption ask respondents if

respondents who feel uneasy to report bribery because they paid bribes may feel at ease discussing other topics they consider less sensible, potentially because they have not been involved in those other socially reprehensible acts. Because reticence is context-specific, using information about sensitive behavior unrelated to corruption may not be a successful way to identify reticence to answer corruption-related questions.

In this article, we propose an alternative method to detect reticent respondents and correct corruption indicators using indirect and direct corruption questions. Since their introduction two decades ago, indirect questions about corruption have been interpreted as reflecting respondent's own experience of corruption despite the literal wording of the question. Recent evidence contradict this loose interpretation. Respondents appear to answer indirect questions literally, estimating the typical corruption behavior of firms like theirs. While they may use their personal experience of corruption when reporting their answers, they most likely also use their perception of their competitors' behavior in their business sector.

Using the literal interpretation of indirect corruption questions allows us to infer the reticence status of a respondent from their response behavior to such questions – a characteristic that is not directly observable. We develop a simple formal model to derive conditions to identify reticent respondents as well as predictions regarding direct and indirect questioning behavior. We exploit the fact that for firm-level survey respondents to genuinely estimate that the average amount of bribes paid by firms similar to theirs is zero (i.e. that not a single firm in their line of business has paid any bribes) is statistically highly improbable, especially in corruption-prone environments. Hence, bribery deniers in indirect questions are identified as probable reticent respondents, deliberately lying to avoid self-incrimination. Such characteristic is then used to identify and correct responses provided by likely reticent respondents to direct corruption questions.

The proposed technique using indirect and direct corruption questions has several advantages over RRQ. Contrary to RRQ which requires implementing additional and non-traditional corruption survey procedures, the proposed technique makes use of currently standard corruption question approaches. The detection of reticent respondents and bias correction technique could hence be used to correct past as well as current corruption surveys. It is efficient in that it uses all information about corruption generated by survey respondents.

Applying our two-step method to a survey of 382 newly created firms in Madagascar, and accounting for the positive correlation between reticence and experience of bribery, we find the unbiased frequency of bribery to be

they have ever made a misstatement on a job application. Others randomized questions relate to personal and business tax fraud.

13 points of percentage greater than the naive estimation of the frequency of bribery not accounting for reticence (42.4 percent and 29.4 percent respectively). It is also 4.9 points of percentage greater than the estimation assuming absence of correlation between reticence and guilt.

This article is organized as follows: section 2 introduces direct and indirect questioning, section 3 presents the model and section 4 the method for correcting indicators for the frequency of bribery from the reticence bias. Section 5 applies this correction to the firm survey carried out in Madagascar. Section 6 concludes.

2 Direct and indirect questioning

The two-step method we propose for detecting reticent respondents and correcting bribery indicators involves both direct and indirect approaches to formulate sensitive questions about bribery. In this section, we briefly present these two approaches and discuss how reticent respondents, despite their non-observable characteristics, can be identified through indirect questioning, and then use this information to correct corruption indicators.

2.1 Direct questioning

Corruption takes different forms. One of the most frequently encountered phenomenon is administrative corruption, through which bribes are requested by public officials or offered by firms to avoid red tape or obtain a service (e.g. permit, license, access to infrastructure)³.

To estimate the extent to which firms are exposed to bribery when interacting with public officials, researchers typically carry out surveys on representative samples of firms, asking owners or managers whether their firm paid any bribes, or alternatively how much their firm paid in bribes, over a certain period of time⁴.

The frequency (magnitude) of bribery can then be estimated by evaluating the share of respondents who report having paid at least one bribe (the average amount of bribes reported). Such question formulation requesting respondents to answer the question based on their own experience of corruption is referred to as “direct questioning”.

The World Bank’s *Enterprise Surveys*, for example, rely on direct questioning to evaluate the proportion of firms who pay bribes to obtain public

³Other forms of corruption are often classified as high-level or “grand” corruption, executive and legislative “state” capture and favoritism (Shaw, 2006).

⁴Linked with respondents, firms, public officials and environmental characteristics, various behavioral, institutional and policy analysis are conducted using these firm-level data (e.g. Fisman and Svensson, 2007; Alm *et al.* 2016).

services such as a water, electricity or telephone connection, a construction permit or a license to import goods or to operate their business.

Some respondents may perceive direct questions to be intrusive or even threatening and, as a consequence, be reluctant to answer them truthfully. To reduce reticence, survey researchers have developed over time various reticence-reducing techniques applied to the design and conduct of surveys on sensitive topics, including the randomized response questioning (RRQ) method discussed above, as well as the use of indirect questioning.

2.2 Indirect questioning

As mentioned, with indirect questions, survey respondents are not asked directly about their own experience of the behavior subject to under-reporting, but indirectly about their perception of the behavior of individuals similar to them. One of the first applications of indirect questioning in corruption surveys was made in the 1998 World Bank's *Ugandan Industrial Enterprise Survey*. Indirect questioning was favored over direct questioning for some of the questions in order to encourage respondents to answer honestly by ensuring that their answers were not self-incriminating (Svensson, 2001).

Like RRQ, indirect questioning gives respondents the opportunity to distance themselves from their answers. Asking some of the questions about bribery indirectly is now standard practice in surveys involving questions regarding corruption. For instance, the World Bank's *Enterprise Surveys* questionnaire contains several questions about bribery that are asked indirectly. One of them asks firm owners and managers to estimate the overall amount of bribes paid annually by firms like theirs. The question in the current version of the questionnaire reads as follows:

We've heard that establishments are sometimes required to make gifts or informal payments to public officials to "get things done" with regard to customs, taxes, licenses, regulations, services etc. On average, what percent of total annual sales, or estimated total annual value, do establishments like this one pay in informal payments or gifts to public officials for this purpose?

(The World Bank, 2016)

Answers to indirect questions – asking respondents to estimate the average amount of bribes paid by firms like theirs – have long been interpreted as if they reflected respondents' own experience of corruption instead of the experience of an average representative firm in their sector, as the question reads. Svensson (2003) for instance looked for correlates between the probability of firms reporting a positive amount of bribes and firm's own characteristics using an indirect question asking Ugandan firms to estimate how much bribes firms similar to theirs "typically pay each year."

However, this interpretation has been questioned. Clarke (2012) emphasized that firms tend to respond literally to an indirectly-phrased question, reporting their estimation of the overall experience of corruption in the sector and not uniquely for their own firm. Using data from a survey of 144 construction firms in Afghanistan, Clarke (2012) showed that firms that are not bidding for government contracts - possibly because they are averse to paying bribes - are more likely than bidding firms to state that firms like theirs make informal payments or gifts to “get things done” with public entities. According to Clarke, this result shows that firms tend to interpret indirect questions literally. Managers answer these questions using their perception of how much firms similar to theirs pay in bribes, and not only their personal experience.

2.3 Using indirect questioning to detect reticence

With direct questioning, there are two types of respondents who report not having personally paid any bribes: candid respondents who truly believe they have not paid bribes, and reticent respondents who purposely lie because they are reluctant to acknowledge to the interviewer having paid bribes. Because reticence is not observable, one cannot estimate from those answers the true proportion of bribe-paying firms.

With indirect questioning, respondents reporting a zero level of bribes among firms similar to theirs could allow generating more information on respondents’ types and may reveal reticence to answer indirect questions. Indeed, despite evidence suggesting that respondents appear to estimate what firms like theirs pay in bribes, reported zero magnitude estimates of bribery is relatively common. For instance, in the firm-level survey examined by Clarke (2012), 62 percent of construction firms in Afghanistan bidding on government contracts reported a zero magnitude of bribery for firms similar to theirs in relation to these contracts. However, reporting that not a single firm similar to the respondent’s firm has paid any bribes may seem particularly unlikely in countries like Afghanistan known to be strongly affected by corruption. Indeed, as the model presented in this paper shows, even under very conservative assumptions, it is highly unlikely that respondents honestly estimate that the average amount of bribes paid by firms similar to theirs is zero.

This response peculiarity was also observed in Madagascar where firms which secured more public contracts identified at high risk of corruption reported more frequently the absence of bribery in public procurement for firms like theirs than firms earning contracts less at risk of corruption (Gauthier and Lesné 2017). This appears to indicate that even if the question does not ask them explicitly about their personal experience of bribery, some respondents choose to deny the existence of bribery altogether for fear

of self-incrimination⁵.

Gauthier and Lesné (2017) also showed that the experience of corruption in public procurement in Madagascar is not related to the likelihood of respondents refusing to participate in a survey about this topic or to answer a sensitive question about the average percentage value of public contracts paid in bribes by firms similar to theirs. Non-response does not appear to be a strategy used by corrupt firms to hide their behavior.

In that respect, respondents with personal experience of corruption appear to act like reticent respondents, knowingly lying to direct questions, but also to indirect questions about corruption when honest answers could lead to the interpretation that they have being involved in corrupt acts themselves. Interestingly, differences in response behavior by respondents to indirect questions could reveal information about their unobservable types. We exploit this information revelation in the approach we propose in this paper to identify reticent respondents.

The formal model proposed in the next section formalizes the behavior of respondents to direct and indirect questions, in particular with regard to non-response and denial of corruption, which is associated to specific respondents' characteristics. Based on the predictions of the model, we believe that all respondents who report the absence of bribery in indirect questioning are likely reticent respondents. We then propose, based on the identification of a set of reticent respondents, a method for correcting indicators for the frequency of bribery using direct questioning.

3 A simple model of response behavior

In this section, we develop a simple formal framework to examine how respondents weight the benefits and costs of truthfully answer direct and indirect questions about corruption.

The decision how to answer bribery questions can be modeled in the framework of the rational choice theory in which survey respondents make use of a cost-benefit analysis to compare the benefits and the costs of providing an honest answer to the question (Krumpal, 2013). Among the benefits for the respondent to provide an honest answer are the satisfaction of respecting social norms of honesty and cooperation, the interest in voicing concerns about the issue of corruption, and the desire to contribute mean-

⁵A more subtle way to hide personal experience of corruption is to underestimate the amount of bribes paid by similar firms while still estimating a strictly positive magnitude of bribery. This strategy is clearly sub-optimal as reticent respondents are bothered even by a slight probability of appearing corrupt. It may however be used by sophisticated respondents who understand that stating that not a single firm similar to theirs has paid any bribes may seem insincere.

ingly to the survey. The costs of an honest answer admitting guilt are in particular the social prejudice of confessing to an interviewer their involvement in corrupt acts, and the threat of legal sanctions or other forms of negative retaliation such as a worsening relationship with public officials. Respondents decide to answer the question honestly if their perceived gains outweigh their perceived costs. Reticent respondents believe they have too little to gain from disclosing socially undesirable behavior to the interviewer with regard to the costs of revealing this information. Because they fear that their answer might be interpreted as an admission of wrongdoing, reticent respondents seek to distance themselves as much as possible from the eventuality of involvement in corrupt acts.

3.1 The model

Consider a survey asking N firm owners and managers about the amount of bribes they paid to public agents when performing an administrative procedure they all had to complete. The survey also asks each of these N firms to evaluate the average amount paid by all surveyed firms, then included. The former question is called “direct”, while the later is referred to as “indirect”.

3.2 Direct questioning

The estimation by respondent $i \in \{1, \dots, N\}$ of the amount of bribes he/she paid, denoted cor_i^D , is defined as follows:

$$cor_i^D = \begin{cases} cor_i^*(X_i) + \varepsilon_i & \text{if } U(X_i, Y = D) \geq 0 \text{ and } \varepsilon_i \geq -cor_i^*(X_i) \\ 0 & \text{if } U(X_i, Y = D) \geq 0 \text{ and } \varepsilon_i < -cor_i^*(X_i) \\ 0 & \text{if } U(X_i, Y = D) < 0 \end{cases} \quad (1)$$

With $cor_i^*(X_i) \geq 0$ the real amount of bribes paid by firm i which is a function of its characteristics X_i . The error term ε_i of the estimation by respondent i of the amount of bribes $cor_i^*(X_i)$ he/she actually paid is a zero-expectation symmetric random variable⁶ of variance $\sigma_{\varepsilon_i}^2$.

$U(X_i, Y = D)$ or U_i^D is a function which evaluates the utility of respondent i to answer the direct question honestly. Firm characteristics X_i affect the benefits and costs of an honest answer, together with a vector Y of features of the survey and question. The vector D is the evaluation of Y when the question is direct. If the benefits for respondent i to answer

⁶The distribution of the error term being centered at zero means that the estimation by respondent i of $cor_i^*(X_i)$ is not systematically biased, either positively or negatively.

honestly overweight the costs of an honest answer, the utility U_i^D is positive or null⁷ and respondent i chooses to answer the direct question candidly.

Those respondents are aware that $cor_i^*(X_i)$, the actual amount of bribes paid cannot be a negative number. If their estimation $cor_i^*(X_i) + \varepsilon_i$ is strictly negative, respondents will adjust it by reporting instead a zero amount.

Reticent respondents to the direct question have a strictly negative utility U_i^D . No matter whether they actually paid bribes or not, reticent respondents will systematically report that they did not paid any bribes⁸.

Finally, when the variance of the error term is higher than a threshold θ_i self-defined by respondent i ($\sigma_{\varepsilon_i}^2 > \theta_i$), cor_i^D is undefined and respondent i refrains from answering the question. The threshold θ_i can be interpreted as the highest level of uncertainty that is acceptable for respondent i to provide an credible estimation of the amount of bribes that he/she paid. From the N surveyed firms, only the subset M of firms whose error variance is below their self-defined threshold agree to answer the direct question. Figure 1 summarizes the process leading to the definition of cor_i^D by respondent i .

3.3 Indirect questioning

The indirect question asks all respondents to estimate the average amount of bribes paid by the whole set of N firms. Concretely, respondent $i \in \{1, \dots, N\}$ estimates how much bribes he/she paid as well as the amount of bribes paid by every other firms in N , and compute an average estimate for those N firms. This estimation, denoted cor_i^I , is defined as follows:

$$cor_i^I = \frac{1}{N} \left(cor_i^D + \sum_{j=1}^{N-1} cor_{ij}^I \right) \quad (2)$$

Firm j is any firm in N other than firm i . cor_{ij}^I is the estimation by respondent i of bribes paid by firm j , defined as follows, $\forall j \neq i \in \{1, \dots, N\}$:

$$cor_{ij}^I = \begin{cases} cor_j^*(X_j) + e_{ij} & \text{if } U(X_i, Y = I) \geq 0 \text{ and } e_{ij} \geq -cor_j^*(X_j) \\ 0 & \text{if } U(X_i, Y = I) \geq 0 \text{ and } e_{ij} < -cor_j^*(X_j) \\ 0 & \text{if } U(X_i, Y = I) < 0 \end{cases} \quad (3)$$

⁷Without loss of generality, respondents answer honestly if their utility is null.

⁸For simplicity, we assume that reticent respondents never admit bribes. A more general model could allow for reticent respondents to sometimes admit bribes, in line with the definition of a reticent respondent proposed by Azfar and Murrell (2009). Reticent respondents who never admit bribes may be seen as displaying the highest level of reticence among reticent respondents. The model presented in this article focuses on identifying respondents belonging to this sub-population of highly reticent respondents.

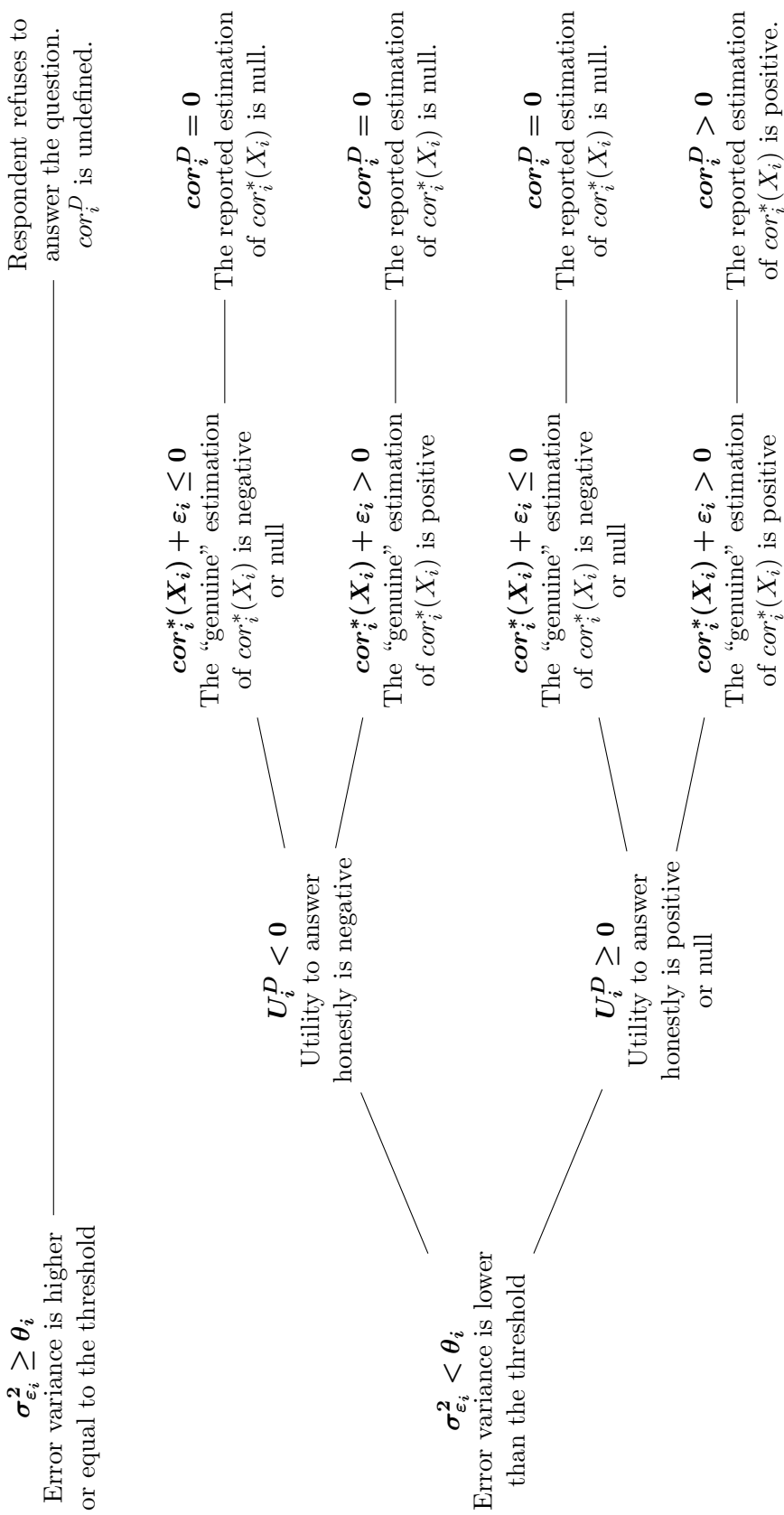


Figure 1: Outcome tree for cor_i^D

The variable $cor_j^*(X_j) \geq 0$ is the amount of bribes really paid by firm j which is a function of its characteristics X_j . $U(X_i, Y = I)$ or $U_i^I < 0$ is the utility function of respondent i to honestly estimate the amount of bribes paid by all firms in N . The arguments of this utility function are a vector of characteristics of firm i and the vector I of survey and question characteristics Y when the question is asked indirectly.

The estimation error by respondent i of the amount of bribes paid by firm j is written e_{ij} . This error term e_{ij} is a zero-expectation symmetric variable with variance $\sigma_{e_{ij}}^2, \forall j \neq i \in \{1, \dots, N\}$.

Whenever respondent i finds a strictly negative value for the amount of bribes paid by any firm in N , he/she corrects it by replacing this negative value with a zero estimate of bribes for that firm.

Respondent i answers the indirect question honestly only if he/she has a positive or null utility U_i^I . If this utility is evaluated to be strictly negative ($U_i^I < 0$), respondent i will always report a zero estimate for the average amount of bribes paid by the N firms.

As discussed in the previous section of this paper, indirect questioning has been introduced in corruption business surveys in order to reduce reticence from firms to reveal sensitive information. Providing an average estimate of bribes paid by firms in their line of business rather than how much they paid may encourage respondents to disclose useful information about the magnitude of bribery without admitting personal involvement in corrupt acts. In reality, it is doubtful whether indirect questioning is actually successful in reducing reticence of survey respondents to answer corruption questions honestly. In our utility framework, this means that the benefit-cost ratio of an honest answer to the indirect question is either higher or equal to that of the direct question. Some respondents who deliberately lie when answering the direct question may answer honestly the indirect question. Formally, $\forall i \in \{1, \dots, N\}: \Pr(U_i^I \geq 0) \geq \Pr(U_i^D \geq 0)$.

Under the assumption of independence of estimations by respondent i of the amount of bribes paid by each firm in N , the variance of his/her “genuine” estimation of the magnitude of bribery is: $\sigma_{\varepsilon_i}^2 + \sum_{j=1}^{N-1} \sigma_{e_{ij}}^2$.

Respondent i agrees to answer the indirect question as long as the variance of his/her estimation is lower than $N\theta_i$. If so, cor_i^I is defined. In the opposite case, $\sigma_{\varepsilon_i}^2 + \sum_{j=1}^{N-1} \sigma_{e_{ij}}^2 \geq N\theta_i$ and respondent i refuses to answer the indirect question. Only the subset L of firms in $\{1, \dots, N\}$ meet this response condition and agree to answer the indirect question.

Finally, we assume that respondent i is less uncertain of the amount of bribes he/she paid than what every other firm in N has paid: $\sigma_{e_{ij}}^2 > \sigma_{\varepsilon_i}^2, \forall j$.

3.4 Predictions of the model

Two main predictions regarding non responses and zero estimates from direct and indirect questioning can be inferred from the model.

3.4.1 Direct questioning has a lower (or equal) non-response rate

Since $\forall j : \sigma_{e_{ij}}^2 > \sigma_{\varepsilon_i}^2$, we know that $\sigma_{\varepsilon_i}^2 + \sum_{j=1}^{N-1} \sigma_{e_{ij}}^2 > N\sigma_{\varepsilon_i}^2$. If the variance of the estimation error for the amount of bribes paid by respondent i is higher than the threshold of the response condition ($\sigma_{\varepsilon_i}^2 > \theta_i$), the variance of the estimation by respondent i of the average amount of bribes paid by the N firms is necessarily higher than $N\theta_i$. In words, respondent i never answers the indirect question if he/she also refuses to answer the direct question.

The reverse scenario is possible: $\sigma_{\varepsilon_i}^2 + \sum_{j=1}^{N-1} \sigma_{e_{ij}}^2 > N\theta_i$ with $\sigma_{\varepsilon_i}^2 \leq \theta_i$. In this case, respondent i agrees to answer the direct question. However, as the level of uncertainty of respondent i with regards to how much bribes firms in N other than his/hers have paid ($\sum_{j=1}^{N-1} \sigma_{e_{ij}}^2$) is relatively large, he/she refrains from answering the indirect question.

Therefore, the probability of respondent i refusing to answer the indirect question is at least as high as the direct question. This translates into a higher (or equal) non-response rate for this question as the set of M firms who agree to answer the direct question is at least as large as the set of L firms who accept to answer the indirect question among the N firms.

3.4.2 Direct questioning has a higher share of zero estimates

With direct questioning, as mentioned above, we could identify two types of respondents reporting a zero estimate of bribery: those who sincerely believe they have not paid any bribes, and those who think they paid a positive amount of bribes but choose to lie deliberately by denying such payment. The probability of respondent i reporting no bribes paid by his/her firm, among the M respondents who agreed to answer the question, is as follows⁹:

$$\Pr(\text{cor}_i^D = 0) = \Pr(U_i^D \geq 0) \Pr(\text{cor}_i^* + \varepsilon_i \leq 0 | U_i^D \geq 0) + \Pr(U_i^D < 0) \quad (4)$$

Similarly, the probability of respondent i to estimate a zero average of bribes for all firms in N , for those L respondents who accepted to answer the indirect question, can be written as follows:

$$\Pr(\text{cor}_i^I = 0) = \Pr(U_i^I \geq 0) \Pr(\text{cor}_i^I = 0 | U_i^I \geq 0) + \Pr(U_i^I < 0) \quad (5)$$

With:

$$\Pr(\text{cor}_i^I = 0 | U_i^I \geq 0) = \Pr(\text{cor}_i^* + \varepsilon_i \leq 0 | U_i^I \geq 0) \times \prod_{j=1}^{N-1} \Pr(\text{cor}_j^* + e_{ij} \leq 0 | U_i^I \geq 0) \quad (6)$$

⁹The arguments X_i and X_j are omitted from subsequent equations for reasons of space.

Since $\Pr(\text{cor}_j^* + e_{ij} \leq 0) \leq 1$ for all j firms¹⁰, $j \neq i \in \{1, \dots, N\}$, the probability that $\text{cor}_i^I = 0$ conditional on $U_i^I \geq 0$ gets closer to 0 as N increases. As equation (6) shows, $\Pr(\text{cor}_i^I = 0 | U_i^I \geq 0)$ approaches zero at an exponential rate with respect to the number of firms N . With $N \rightarrow \infty$:

$$\Pr(\text{cor}_i^I = 0) = \Pr(U_i^I < 0) \quad (7)$$

As N grows large, it is more and more likely that respondents who estimate the magnitude of corruption in N to be 0 (i.e. that respondents estimate that no firm in N has paid any bribes) are reticent.

To illustrate how likely an honest respondent would estimate a zero magnitude of bribery in N , let us first consider the situation in which not a single firm in N has actually paid any bribes.

Assuming all error terms $\{\varepsilon_i, e_{i1}, \dots, e_{i(N-1)}\}$ in (6) are continuous random variables independent of the level of utility U_i^I of respondent i to answer the indirect question, we know that $\Pr(\varepsilon_i \leq 0 | U_i^I \geq 0) = \Pr(\varepsilon_i \leq 0) = \Pr(\varepsilon_i < 0)$ and $\Pr(e_{ij} \leq 0 | U_i^I \geq 0) = \Pr(e_{ij} \leq 0) = \Pr(e_{ij} < 0), \forall j \neq i \in \{1, \dots, N\}$. We can thus rewrite equation (6) as:

$$\Pr(\text{cor}_i^I = 0 | U_i^I \geq 0) = \Pr(\varepsilon_i < 0) \times \prod_{j=1}^{N-1} \Pr(e_{ij} < 0) \quad (8)$$

Using the symmetry property of error terms, $\Pr(\text{cor}_i^I = 0 | U_i^I \geq 0) = (0.5)^N$.

Figure 2 presents the respondent's probability of estimating a zero magnitude of bribery. This probability reaches zero at an exponential rate with respect to the number of firms N . When $N = 10$, the probability of respondent i to estimate a zero magnitude of bribery in N if he/she answers the question honestly is $(0.5)^{10} \approx 0.1\%$. With $N = 20$, this probability is lower than 0.0001%.

However, if errors are discrete instead of continuous variables with a strictly positive probability mass at 0 (estimation with no error), the probability of an honest answer to the indirect question being null is:

$$\Pr(\text{cor}_i^I = 0 | U_i^I \geq 0) = \left(\frac{1 + \gamma_{\varepsilon_i}}{2}\right) \times \prod_{j=1}^{N-1} \left(\frac{1 + \gamma_{e_{ij}}}{2}\right) \quad (9)$$

With γ_{ε_i} and $\gamma_{e_{ij}}$ the probability mass at 0 of ε_i and e_{ij} , respectively. With a 50 percent chance of respondent i correctly estimating the amount

¹⁰ $\Pr(e_{ij} > 0) > 0, \forall j$ as $E(e_{ij}) = 0$ and $\sigma_{e_{ij}}^2 > 0$ since $\sigma_{e_{ij}}^2 > \sigma_{\varepsilon_i}^2$. With $\text{cor}_j^* \geq 0$, $\Pr(\text{cor}_j^* + e_{ij} > 0) > 0$ and $\Pr(\text{cor}_j^* + e_{ij} \leq 0) \leq 1, \forall j$.

of bribes paid by all firms in N , the probability of honestly estimating that no bribes were paid by any firms in N is $(0.75)^N$. This probability is 5.6% with $N = 10$ and 0.3% if $N = 20$ (Figure 2).

Clearly, the probability of respondent i estimating a zero average of bribes for all firms in N decreases faster if at least one firm has paid bribes. The speed at which this probability tends to zero then depends on the proportion of corrupt firms (i.e. the frequency of bribery in N), the amount of bribes paid by corrupt firms as well as the variance of the error terms.

Even under conservative assumptions regarding the frequency of corruption in N as well as the proportion of respondents who make no error in estimating bribes, the probability of a zero estimate for the magnitude of bribery in N among honest respondents quickly reaches very low values as N increases. With N large, one can safely consider that all respondents providing a zero answer to the indirect question are reticent respondents.

In summary, while respondents providing a zero estimate of bribes to the direct question are either reticent or actually believe they did not pay any bribes, a prediction of the model is that respondents to indirect questioning reporting a zero average estimate of bribes paid by all firms could be assumed to be reticent respondents, providing the number of firms assessed is sufficiently large. Furthermore, given that non-response is driven by the level of uncertainty of respondents with regard to their estimates, and assuming that respondents are less uncertain of their own experience of bribery than of that of their competitors, non-response is likely higher for indirect questions than for direct questions.

4 An approach for correcting the frequency of bribery from the measurement bias caused by reticence

Having examined the response behavior to direct and indirect questioning, we propose in this section a method to correct bribery indicators from the measurement bias caused by reticence.

Following Clarke *et al.* (2015) or Azfar and Murrell (2009), we focus on the frequency of bribery rather than the average amount of bribes paid given the sensitivity of estimates of the magnitude of bribery to the phrasing of the survey questions¹¹.

¹¹Clarke (2009) and Lesné (2017) independently showed that indicators for the amount of bribes paid by firms (i.e. the magnitude of bribery) are considerably higher when respondents are asked to report their estimates as a percentage of their sales rather than in monetary terms.

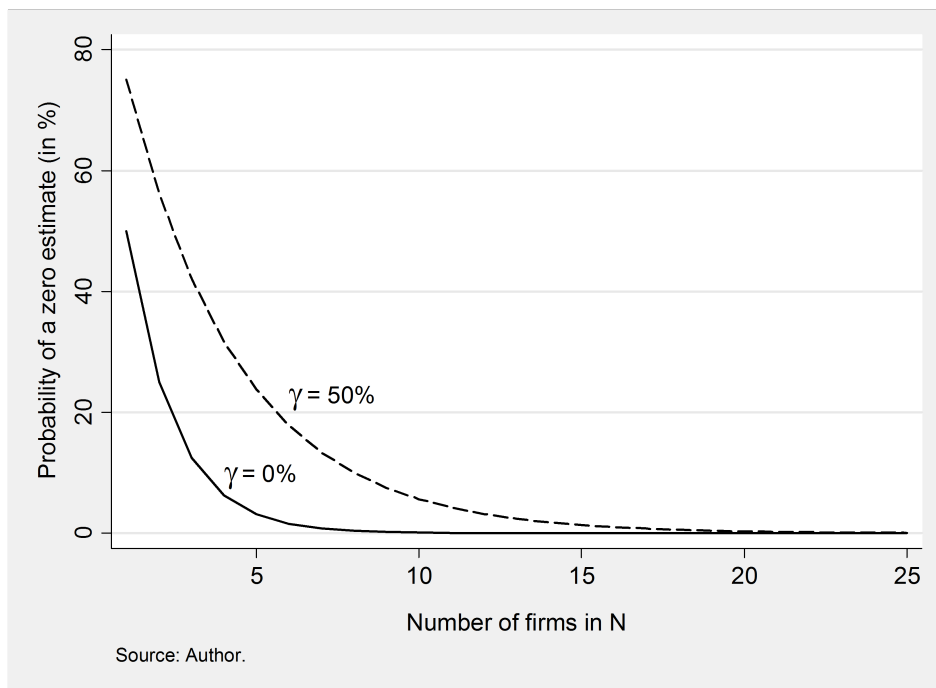


Figure 2: Respondent's probability to estimate a zero magnitude of bribery in N , by number of firms in N and by probability γ of the respondent making no error in estimating this magnitude (with the assumption that no firm in N actually paid any bribes).

4.1 A probabilistic indicator for the frequency of bribery

The frequency of bribery is defined as the proportion of individuals in a designated population who paid bribes over a certain period of time. This frequency can be estimated from a given survey sample by computing the number of survey respondents who reported having paid bribes, and dividing this number by the total number of surveyees except for those who did not provide the requested information. This computation may however be affected by measurement bias due to the presence of reticent respondents in the population and sample.

Using the framework introduced in the previous section, we define a probabilistic indicator for the frequency of bribery on a set of N firms:

$$Freq_{cor} = \frac{\Pr(cor_i^* > 0) + \sum_{j=1}^{N-1} \Pr(cor_j^* > 0)}{N} \quad (10)$$

This indicator is estimated using responses to the direct question from the subset of $M \in \{1, \dots, N\}$ firms who answered this question, including firm i .

$$\widehat{Freq}_{cor} = \frac{\Pr(cor_i^D > 0) + \sum_{j=1}^{M-1} \Pr(cor_j^D > 0)}{M} \quad (11)$$

With cor_j^D the estimation by respondent $j \neq i \in \{1, \dots, M\}$ of the amount of bribes he/she paid. Surveyees who refused to answer the direct question and, as such, were not included in the set $\{1, \dots, M\} \in \{1, \dots, N\}$, are left out as no information is available about their experience of bribery.

4.2 Identifying the reticence bias

From (1), we know that:

$$\Pr(cor_i^D > 0) = \Pr(U_i^D \geq 0) \Pr(cor_i^* + \varepsilon_i > 0 | U_i^D \geq 0) \quad (12)$$

If reticence were non-existent, cor_i^D would be a suitable estimator¹² of cor_i^* since $E(\varepsilon_i) = 0$. With the decision process regarding non-response being independent from the amount of bribes respondents paid ($\theta_i \perp\!\!\!\perp cor_i^*, \forall i \in \{1, \dots, N\}$), estimating the frequency of bribery would straightforwardly rely on data collected from the subset of M respondents and equation (11).

¹²We must note that the correction operated by respondent i on cor_i^D when $\varepsilon_i < -cor_i^*$ generates a slight upward bias on cor_i^D , which is all the more important as $\sigma_{\varepsilon_i}^2$ is large.

Since $\Pr(U_i^D \geq 0) + \Pr(U_i^D < 0) = 1$, the “genuine” estimation by respondent i of the amount of bribes he/she paid can be written as follows:

$$\begin{aligned} \Pr(\text{cor}_i^* + \varepsilon_i > 0) = & \\ & \Pr(U_i^D \geq 0) \Pr(\text{cor}_i^* + \varepsilon_i > 0 | U_i^D \geq 0) + \\ & \Pr(U_i^D < 0) \Pr(\text{cor}_i^* + \varepsilon_i > 0 | U_i^D < 0) \quad (13) \end{aligned}$$

Because $\Pr(U_i^D < 0) > 0$, we have: $\Pr(\text{cor}_i^* + \varepsilon_i > 0) > \Pr(\text{cor}_i^D > 0)$. The positive probability of having reticent respondents in the survey sample generates a downward bias in our estimator \widehat{Freq}_{cor} .

The size of the reticence bias for respondent i is:

$$Bias_i(\widehat{Freq}_{cor}) = \Pr(U_i^D < 0) \Pr(\text{cor}_i^* + \varepsilon_i > 0 | U_i^D < 0) \quad (14)$$

Equation (14) provides two useful insights for correcting the reticence bias. The first is that only reticent respondents who believe they paid a positive amount of bribes are problematic. As discussed in the introduction, reticent respondents tend to avoid admitting involvement in corrupt acts, but are not always involved in such acts. That said, as shown in Gauthier and Lesné (2017), firms involved in corruption may be more likely than honest firms to behave as reticent respondents in business surveys. If so, and this is the second insight of equation (14), it is likely that $\Pr(\text{cor}_i^* + \varepsilon_i > 0 | U_i^D < 0) > \Pr(\text{cor}_i^* + \varepsilon_i > 0 | U_i^D \geq 0)$. Reticence is most plausibly positively correlated with bribery: reticent respondents possibly have more personal experience of bribery, on average, than candid respondents.

4.3 Correcting the reticence bias

By systematically denying payment of bribes for fear of self-incrimination, reticent respondents contribute to lower the estimation for the frequency of bribery obtained from respondents’ answers to the direct question. Correcting the reticence bias for the frequency of bribery requires identifying which respondents are reticent and editing their answers to the direct question so that they reflect their true experience of bribery. The two-step method is hence as follows:

Step 1: identifying reticent respondents

Reticence is a hidden characteristic of survey respondents. Because reticence is not random but a result of the cost-benefit analysis of formulating an honest answer, it can however relate to observable characteristics of respondents such as their age and level of education as well as the size and profitability of their firm. The key element to determining which respondents are reticent in a given survey sample is their response behavior with

regards to the indirect question. As long as respondents estimate the magnitude of bribery based on a sufficiently large number of firms, only reticent respondents estimate that bribes are never paid by firms similar to theirs.

A first variant of the detection method of reticent respondents is to split the survey sample into two randomly defined groups of respondents. The direct question is then asked only to one of these two groups, while the indirect question is only asked to respondents belonging to the other group. All respondents to the indirect question who estimate a zero magnitude of bribery are labeled reticent, in accordance with the findings of the model of response behavior.

Regressing the response status (zero versus positive estimation of the magnitude of bribery) on observable characteristics of respondents to the indirect question informs about the typical profile of a reticent respondent. Respondents to the direct question who report that they did not pay any bribes whose characteristics match the most closely the profile of a reticent respondent are consequently identified as reticent. As the two groups of respondents to the direct and indirect questions are generated randomly, the proportion of reticent respondents to the indirect question is used to determine the number of reticent respondents to the direct question¹³.

Another, simpler variant of the detection method of reticent respondents, is to ask the two questions, indirect and direct, consecutively to all survey respondents. In this variant, the indirect question is asked first to detect reticent respondents. The information is then immediately used to identify reticent respondents from candid respondents among respondents who reported not having paid bribes in the following (direct) question.

Step 2: correcting the frequency of bribery indicator

Once reticent respondents to the direct question have been identified, the next step in correcting the frequency of bribery indicator from the reticence bias is to eliminate the unwanted influence of answers from reticent respondents in the survey sample. One option followed by Azfar and Murrell (2009) is to remove all respondents identified as reticent from the computation of the indicator. By doing so, they assumed that reticent respondents are not more likely than candid respondents to pay bribes. However, as emphasized by Clarke (2012), treating reticent respondents as if they were not significantly different from honest respondents regarding their personal experience of corruption may lead to misleading results. Indeed, as shown in Gauthier

¹³It is possible, although not obvious, that reticence is less acute for indirect questioning than for direct questioning. This means that some respondents who are reticent to answer the direct question would nevertheless accept the answer honestly to the indirect question. If this is the case, the number of reticent respondents to the direct question is underestimated with our method, all the more as indirect questioning is successful in reducing reticence.

and Lesné (2017), firms with more experience of corruption tend to act more frequently like reticent respondents than virtuous firms. It is then plausible that personal experience of bribery and reticence are in fact positively correlated. Simply removing reticent respondents from the computation of the indicator for the frequency of bribery - or similarly assigning to reticent respondents the same average probability of having paid bribes as the group of candid respondents - will not fully purge the indicator from the reticence bias.

A more adequate option followed in this proposed method is to use answers of candid respondents to the direct question to learn about the firm characteristics associated with the probability of paying bribes. With this explanatory model of bribe payment as a function of firm characteristics, it is then possible to define the probability of bribe payment for each respondent identified as reticent based on their own characteristics. We assume that reticent respondents whose probability of having paid bribes is equal to or higher than 50 percent, are classified as bribe payers.

5 An application to Madagascar

In this section, we test the predictions of the formal model and apply the proposed two-step method for correcting bribery indicator from the reticence bias to the case of a firm-level survey conducted in Madagascar. We first present the context of the corruption survey and then assess the extent of measurement bias associated with reticent respondents.

5.1 The randomized corruption survey

A survey was conducted among 382 Malagasy firms by the anti-corruption NGO Transparency International in March 2016. The objective of the survey was to identify corruption-related constraints associated with establishing a business in Antananarivo, Madagascar’s capital city. The survey was designed and supervised by the authors of this paper.

Following a series of questions on the nature and the costs of creating their business, firm owners and managers were asked to estimate the total amount spend to formalize their business, summing up official costs of the procedure with any additional gifts or informal payments. To assess the extent of bribe payments, respondents were also asked either a direct or an indirect question on bribe payments. Under the direct formulation, respondents had to estimate the total amount of “gifts and other informal payments” that their firm paid to complete their start-up procedures. Under the indirect formulation, respondents had to estimate how much firms similar to theirs need to pay, on average, to complete their own procedures.

The survey sample was randomly split between those assigned the direct or indirect question. The sub-samples of the two versions of the ques-

tionnaire were equal, with 191 respondents each¹⁴.

Table 1 shows the characteristics of the respondents to the direct and indirect versions of the bribery question. We observe that the average characteristics of the two groups are similar. About two-thirds (67.8%) of firms in the sample were registered in the first borough, which is the largest borough of the city. Women make up slightly less than half of the sample (44.2%). Respondents are predominantly middle-aged (62.3% between 30 and 49 years old) and mostly have a higher education (53.7%). Sole proprietorship is by far the most common legal form adopted by new firms in Antananarivo (81.4%)¹⁵. We also note that about half of surveyed firms operate in retail trade (43.7%), services being the second most important sector (27.2%), followed by transport (15.2%).

Given that the observable average characteristics of the firms which were randomly allocated the two versions of the questionnaire are not statistically different at conventional levels, we can straightforwardly evaluate how the formulation of the bribery question affects response behavior.

5.2 Testing the model's predictions

The model predicts that the non-response rate is higher for the indirect formulation compared to the direct formulation, and that the rate of zero estimates of bribery is lower for the indirect formulation than for the direct formulation. These two model's predictions are verified in the Madagascar survey.

5.2.1 Non-response

As observed in Table 2, 38.2 percent of respondents did not provide an estimate for the average amount of bribes paid by firms similar to theirs when formally registering their activity (indirect question). In comparison, the non-response rate to the direct question, which asked respondents to evaluate how much their firm paid to complete their registration process, is about twice as low (19.9%). The difference in proportions of non-respondents to the two questions is statistically significant at 1 percent (Table 2). Inter-

¹⁴As explained earlier, a suitable variant to this random allocation of the versions of the question between two groups would be for all respondents to first answer the indirect version of the question, and then answer the direct question version immediately after. In Madagascar, a randomized experiment was chosen in order also to test the implications of the model introduced in this paper. A randomized experiment provides a more adequate research framework to assess the influence of the question format on the response rate and on the proportion of zero estimates in the sample by excluding the possible interference of the indirect question on later answers to the direct question.

¹⁵Other legal forms are limited liability company (*Société à responsabilité limitée* - SARL) and limited company (*Société anonyme* - SA).

viewers have classified the reasons given by respondents for failing to answer the questions into refusal to answer (11.7%) and lack of sufficient knowledge (82.9%)¹⁶. Interestingly, the difference in the non-response rates to the two questions is statistically significant for the lack of sufficient knowledge, but not for the refusal to answer. This confirms that reticence to answer the two questions is likely equivalent and that the higher non-response rate for the indirect question presumably results from the complexity for respondents to accurately estimate what other firms pay, as assumed by the model.

5.2.2 Zero estimates

A fundamental prediction of the response behavior model is that the proportion of respondents to answer zero to the bribery question is lower for the indirect formulation compared to the direct formulation. Including respondents who did not provide an answer to the question, the proportion of firm owners or managers in Madagascar reporting that they did not pay any bribe at the time of the creation of their business (direct questioning) is 51.3 percent (see Table 2). Respondents who had to estimate instead the average amount of bribes paid by firms similar to theirs to set-up their business (indirect questioning) are “only” 26.2 percent to answer a null amount. As with the non-response rates, the difference in the zero response rates between the two formulations of the bribery question is statistically significant at 1 percent (Table 2). Excluding non-response, the proportion of respondents to estimate a strictly positive amount of bribes is 57.6 percent for the indirect question and 36.0 percent for the direct question¹⁷, a result compatible with the predicted implications of the model.

5.3 Correcting the reticence bias

Having tested the main predictions of the model, we now apply the two-step method for correcting frequency indicators of bribery of the reticence bias.

Step 1: identifying reticent respondents

As mentioned, respondents estimating in a direct question that they did not pay any bribes when completing the procedure to register their business may do so for two reasons: because they really believe they did not pay any bribes, or because they think they did pay bribes but lie. We have learned from the model that we can safely consider that all respondents to

¹⁶In 6 out of 111 instances (5.4%) of non-response to the direct and indirect questions (3 cases each), surveyors failed to note the reason given by respondents for non-response or could not classify the reason as either refusal to answer or lack of sufficient knowledge.

¹⁷The difference in strictly positive estimates of bribes is also significant at 1 percent.

an indirect question who estimated that not a single firm similar to theirs paid any bribes (or alternatively that the average amount of bribes paid by these firms is null) are reticent. Because the sub-sample of respondents to the indirect question can be split into two groups of reticent and possibly candid respondents¹⁸, we can estimate using a regression model how respondents' characteristics affect their probability of being reticent. We can then predict which respondents to the direct question are likely reticent based on their own characteristics¹⁹.

Since being a reticent or a possibly candid respondent is a binary outcome, a limited dependent variable model - in this case a probit model - is appropriate. We use all descriptive variables described in the previous section as explanatory variables to explain the probability of respondents to the indirect question to provide a strictly positive estimate of the magnitude of bribery. Column (1) of Table 3 reports the results of this probit model estimated with maximum likelihood. The probability of being a reticent respondent appears to be lower for respondents with a higher education level, and also lower for sole proprietorship firms. The McFadden R^2 of this regression run on 115 observations²⁰ is 0.164.

We then use this empirical model to identify which respondents who reported that they did not pay bribes to the direct question share the most commonalities with reticent respondents to the indirect question. The proportion of reticent respondents to the indirect question (42.4%) - those who estimated that no firm similar to theirs paid any bribes - is assumed to be the same as the proportion of reticent respondents to the direct question²¹.

Among the 93 respondents to the direct question who said they did not pay bribes for which regression data are available²², 40 respondents (43.0%) with the highest probability of being reticent according to the empirical model are labeled reticent and the remaining 53 (57.0%) as possibly candid.

¹⁸Like Azfar and Murrell (2009) and others, we use the term "possibly candid" rather than "candid" to label respondents who are not identified as reticent to emphasize that sophisticated reticent respondents who estimate a positive magnitude of bribery to the indirect question are unidentified by the detection technique we propose. Similarly, if indirect questioning reduces reticence with respect to direct questioning, some reticent respondents to the direct question may answer honestly the indirect question, and for this reason be wrongly classified as candid respondents.

¹⁹If the questionnaire includes both direct and indirect questions, this step is unnecessary as all respondents must reveal their response behavior to the indirect question.

²⁰From the 191 respondents who were asked the indirect question, 73 chose not to answer. In addition, age and education data were missing for three respondents to this version of the questionnaire, making a total of 115 observations for this regression.

²¹As mentioned earlier, if indirect questioning is successful in reducing reticence, the proportion of reticent respondents to the direct question is underestimated.

²²Five respondents to the direct question who estimated they did not pay bribes did not provide information about either their age or their level of education. Consequently, the probability of being reticent could not be computed for these five respondents.

Table 1: Characteristics of respondents and firms, by questions (in %)

Proportional share (in %)	Direct questioning	Indirect questioning	Test of difference ⁽¹⁾	All sample	Number of observations
Proportion of firms registered in borough I	69.1	66.5	0.584	67.8	382
Proportion of female respondents	44.5	44.0	0.918	44.2	382
Proportion of respondents aged 20 to 29	21.7	19.0	0.839 ⁽²⁾	20.3	379
Proportion of respondents aged 30 to 39	31.8	34.2		33.0	
Proportion of respondents aged 40 to 49	30.2	28.4		29.3	
Proportion of respondents aged 50 or more	16.4	18.4		17.4	
Proportion of respondents with higher education	57.0	50.5	0.210	53.7	374
Proportion of sole proprietorship firms	79.6	83.3	0.357	81.4	382
Proportion of firms in retail trade	41.9	45.6	0.702 ⁽³⁾	43.7	382
Proportion of firms in wholesale trade	4.7	3.7		4.2	
Proportion of firms in services	27.8	26.7		27.2	
Proportion of firms in transport	14.1	16.2		15.2	
Proportion of firms in other sectors	11.5	7.9		9.7	

Notes: (1) Test of differences are based on Pearson's χ^2 statistic; p-values are reported. (2) Test of differences for the distribution of age groups in the two versions of the questionnaire. (3) Test of differences for the distribution of business sector categories in the two versions of the questionnaire.

Table 2: Non-response rates and proportions of zero estimates, by questions

Indicator (in %)	Direct questioning	Indirect questioning	Test of difference ⁽¹⁾	All sample
Proportion of non-response	19.9	38.2	< 0.001	29.1
<i>Lack of knowledge</i>	14.7	33.5	< 0.001	24.1
<i>Refusal to answer</i>	3.7	3.1	0.778	3.4
<i>Undefined reason</i>	1.6	1.6	-	1.6
Proportion of zero estimates	51.3	26.2	< 0.001	38.7

Notes: (1) Test of differences are based on Pearson's χ^2 statistic; p-values are reported.

Table 3: Empirical model for detecting reticent respondents and correcting the indicator for the frequency of bribery

VARIABLES	(1) Detection	(2) Correction
borough (first)	-0.394 (0.302)	-0.129 (0.338)
respondent gender: female	0.298 (0.271)	0.0903 (0.278)
respondent age: 30 to 39	-0.304 (0.361)	-0.826** (0.406)
respondent age: 40 to 49	-0.384 (0.406)	-0.497 (0.394)
respondent age: 50 or more	-0.114 (0.427)	-0.761* (0.455)
respondent education: higher	-0.972*** (0.330)	-0.420 (0.312)
firm status: proprietorship	-0.889** (0.419)	-0.999** (0.395)
firm sector: retail trade	0.593 (0.546)	-0.505 (0.410)
firm sector: wholesale trade	1.138 (0.742)	0.297 (0.807)
firm sector: services	0.910 (0.579)	0.0974 (0.449)
firm sector: transport	-0.624 (0.664)	0.113 (0.463)
Constant	0.914 (0.791)	1.899*** (0.611)
McFadden R^2	0.164	0.102
Observations	115	107

Coefficients are reported
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Step 2: correcting the frequency of bribery indicator

The naive estimation of the frequency of bribery, not accounting for the measurement bias caused by reticence, would account for the number of respondents reporting having paid a positive amount of bribes divided by the total number of respondents who were asked the direct question, including those who did not answer the question²³. In the case of the Madagascar survey, this frequency of bribery corresponds to 29.4 percent of respondents.

Having identified reticent respondents to the direct question, we can produce an improved indicator for the frequency of bribery corrected from the bias due to the presence of reticent respondents in the survey sample.

Removing reticent respondents from the computation of the indicator, as discussed earlier, was the option chosen by Azfar and Murrell (2009), Clausen *et al.* (2010), and Clarke *et al.* (2015) to correct their indicators for the frequency of bribery. Using this approach for computing the bribery indicator on the Madagascar survey using only answers from possibly candid respondents gives a frequency of bribery of 37.5 percent. This new estimate is 27.6 percent higher than the naive estimate which is very close to Azfar and Murrell (2009)'s results for the Romanian firm survey²⁴. However, as previously discussed, their approach assumes that reticence and personal experience of bribery are uncorrelated, which is unlikely. If reticence and bribery are positively correlated, simply removing respondents identified as reticent from the definition of the frequency of bribery is insufficient to obtain unbiased estimates.

A more effective approach is to determine the probability that reticent respondents did pay bribes, and compute an indicator for the frequency of bribery using probabilities of bribe payment from all respondents, as proposed in our approach.

To do so, we first identify firm characteristics associated with bribe payment by regressing on the sub-sample of possibly candid respondents to the direct question a dummy variable (equal to 1 if the respondent admitted paying bribes and 0 otherwise) using a probit model with a set of firm characteristics as independent variables²⁵.

The results of this regression are presented in Table 3 column 2. We observe that younger respondents appear to pay bribes in a significantly larger

²³For the sake of comparability, we exclude from all estimations the five respondents whom we were unable to determine whether they are reticent or possibly candid.

²⁴Their naive estimate for the frequency of bribery was 14.9 percent. Assuming 35 percent of reticent respondents, their estimate of bribery using only reports from possibly candid respondents was 19.9 percent. The frequency of bribery excluding reticent respondents is 33.6 percent higher than the naive estimate (Azfar and Murrell, 2009).

²⁵The firm characteristics used in this empirical model are exactly the same as the ones used for the detection model of reticent respondents, but they need not be.

proportion than older respondents. Sole proprietorships are less affected by bribe payment than firms registered under another legal status (limited liability company or limited company). Dummies for the firm's business sector are collectively insignificant²⁶, meaning that new firms from all sectors are affected by bribery in equal proportions.

In a final stage, we predict the probability of reticent respondents to the direct question to have paid bribes according to their characteristics and the outcomes of the empirical model for bribe payment. The frequency of bribery among reticent respondents is estimated at 60 percent, as compared to 48.7 percent for the possibly candid respondents (excluding non-response). This confirms that reticent respondents have a higher average probability of paying bribes and that the assumption of absence of correlation between reticence and experience of bribery probably does not hold.

With the imputation of predicted probabilities of bribe payment for reticent respondents, the frequency of bribery is estimated to be 42.4 percent. This corrected frequency of bribery taking into account the positive correlation between reticence and experience of bribery is 44.2 percent higher than the naive estimation, or 13.1 percent higher than the estimation of the frequency of bribery obtained from only possibly candid respondents.

It should be noted that this estimator is a lower bound for the frequency of bribery in the population of firms as it assumes that reticence is as likely for the indirect question as it is for the direct question. It also does not account for the possibility that sophisticated reticent respondents report a positive - although underestimated - magnitude of bribery to the indirect question as they understand that estimating that no firm similar to theirs has paid any bribes is unrealistic, and therefore suspicious.

6 Conclusions

In this article, we propose an innovative method for detecting reticent respondents in firm-level surveys and correcting indicators of corruption from the bias associated with the presence of reticent respondents. This method uses indirect and direct questions on bribery payments. An indirect question is used to identify reticent respondents, which makes it possible to adjust responses to a direct question assessing bribery activities. Applying this method to a survey of newly created firms in Madagascar, and assuming first that reticence and bribery are uncorrelated, we find that the frequency of bribery is underestimated by 27.6 percent. Removing the assumption of absence of correlation between reticence and experience of bribery, the unbiased frequency of bribery is estimated at 42.4 percent, which is 44.2 percent

²⁶The p-value of the F-test is 0.286.

higher than the naive estimation that does not account for reticence.

Our proposed technique using indirect and direct corruption questions has several advantages over previous methods, especially RRQ. Contrary to RRQ which requires implementing non-traditional survey procedures, the proposed technique makes use of currently standard corruption question approaches. The reticent detection and bias correction technique could hence be used to correct past as well as current corruption surveys. It is efficient in that it uses all information about corruption generated by respondents.

The estimator we propose to correct the measurement bias due to reticent respondents is a lower bound estimator as it assumes that reticence is as likely for the indirect question as it is for the direct question. Furthermore, it does not account for the possibility of sophisticated reticent respondents reporting a positive but underestimated magnitude of bribery. In future research, these issues should be further investigated to better understand the relationship between reticence and corruption and mechanisms to generate unbiased corruption estimates.

References

- Alm, J., Martinez-Vazquez, J., and McClellan, C. (2016). Corruption and firm tax evasion. *Journal of Economic Behavior & Organization*, 124:146–163.
- Azfar, O. and Murrell, P. (2009). Identifying reticent respondents: assessing the quality of survey data on corruption and values. *Economic Development and Cultural Change*, 57(2):387–411.
- Clarke, G. R. (2011). How petty is petty corruption? Evidence from firm surveys in Africa. *World Development*, 39(7):1122–1132.
- Clarke, G. R. (2012). What do managers mean when they say ‘firms like theirs’ pay bribes? *International Journal of Economics and Finance*, 4(10):161.
- Clarke, G. R., Friesenbichler, K. S., and Wong, M. (2015). Do indirect questions reduce lying about corruption? evidence from a quasi-field experiment. *Comparative Economic Studies*, 57(1):103–135.
- Clausen, B., Kraay, A., and Murrell, P. (2010). Does respondent reticence affect the results of corruption surveys? evidence from the world bank enterprise survey for nigeria. The World Bank Policy Research Working Paper 5415.
- Fisher, R. J. and Tellis, G. J. (1998). Removing social desirability bias with indirect questioning: is the cure worse than the disease? *ACR North American Advances*.

- Fisman, R. and Svensson, J. (2007). Are corruption and taxation really harmful to growth? firm level evidence. *Journal of Development Economics*, 83(1):63–75.
- Gauthier, B. and Lesné, F. (2017). Is silence an admission of guilt? Working paper FERDI.
- Jensen, N. M., Li, Q., and Rahman, A. (2010). Understanding corruption and firm responses in cross-national firm-level surveys. *Journal of International Business Studies*, 41(9):1481–1504.
- Kraay, A. and Murrell, P. (2016). Misunderestimating corruption. *Review of Economics and Statistics*, 98(3):455–466.
- Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: a literature review. *Quality & Quantity*, 47(4):2025–2047.
- Lensvelt-Mulders, G. J., Hox, J. J., Van der Heijden, P. G., and Maas, C. J. (2005). Meta-analysis of randomized response research: Thirty-five years of validation. *Sociological Methods & Research*, 33(3):319–348.
- Lesné, F. (2017). The same, only different: estimating the magnitude of bribery in business surveys. Working paper FERDI 198.
- Schaeffer, N. C. (2000). Asking questions about threatening topics: A selective overview. *The science of self-report: Implications for research and practice*, pages 105–121.
- Shah, A. (2006). Corruption and decentralized public governance. The World Bank Policy Research Working Paper 3824.
- Svensson, J. (2001). The cost of doing business: Firms’ experience with corruption. *Uganda’s Recovery: The Role of Farms, Firms, and Government*, pages 319–341.
- Svensson, J. (2003). Who must pay bribes and how much? evidence from a cross section of firms. *The Quarterly Journal of Economics*, 118(1):207–230.
- The World Bank (2016). Enterprise surveys: Indicator descriptions.
- Tourangeau, R. and Yan, T. (2007). Sensitive questions in surveys. *Psychological bulletin*, 133(5):859.
- Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60(309):63–69.

Appendix: The Madagascar firm-level survey

The survey asked firm owners and managers about the administrative procedure they went through to formally register their company and start up their business activities, and in particular whether bribes were requested by public officials at each step of this procedure.

The typical process of setting up a business in Madagascar requires from the firm owner some evidence of their identity and a known physical address for their business, the registration of the company at the Madagascar National Statistics Institute (*Institut National de la Statistique* - INSTAT) as well an advance payment on the business income tax for the current fiscal year. Additional documents and formalities may be required depending on the firm's main activity. For example, a restaurant serving alcohol must obtain a specific liquor license issued by the Ministry of the Interior.

The survey sample was drawn from databases of companies created in 2015 obtained from tax centers located in the first and fourth *arrondissement* (boroughs) of the city. Centralized data for all newly created Malagasy firms were not publicly available. This required approaching each Antananarivo's *arrondissement* individually to obtain their data. These two databases contain information about 1,747 firms formally registered during the year 2015 in those two boroughs, including their postal address and phone number.

A team of twelve interviewers first attempted to contact all companies by phone to speak with their owner in order to propose them to take part in the survey. They introduced themselves as surveyors recruited by the NGO Transparency International to carry out a survey investigating conditions for business creation in Madagascar, without specifying at this time that a large proportion of questions were focused on payment of bribes. The survey was conducted with 382 business representatives who were successfully contacted and who accepted to participate in the survey. In 65 cases out of 382 (17%), surveyors could not interview the firm owner but either the manager or a family member of the owner who confirmed having the required knowledge to answer questions about the creation process of the firm. Those respondents do not behave differently from firm owners in terms of how they answer questions about bribery. Consequently, the results we present in this article are for all respondent types.

The survey questionnaire consisted of about 50 questions asking respondents about the requested documents and mandatory administrative formalities at each stage of the formal establishment of their firm, including how much they had to pay for these documents and formalities. In addition, the questionnaire incorporated a series of questions about characteristics of respondents and their firms, as well as questions on respondents' perception of the business environment and the effectiveness of the administration.

“Sur quoi la fondera-t-il l'économie du monde qu'il veut gouverner? Sera-ce sur le caprice de chaque particulier? Quelle confusion! Sera-ce sur la justice? Il l'ignore.”

Pascal



Created in 2003 , the **Fondation pour les études et recherches sur le développement international** aims to promote a fuller understanding of international economic development and the factors that influence it.

 **Contact**

www.ferdi.fr

contact@ferdi.fr

+33 (0)4 73 17 75 30