The same, only different: estimating the magnitude of bribery in business surveys

Frédéric LESNÉ

➜ FRÉDÉRIC LESNÉ is a PhD candidate in Development Economics , CERDI - Université Clermont Auvergne.

Abstract

As first shown in Clarke (2011), Word Bank's Enterprise Surveys data suggest that the magnitude of bribery estimated by firm owners and managers in Africa is considerably higher when estimates are formulated as a percentage of turnover rather than in monetary value. This working paper confirms these findings with a randomized experiment carried out in Madagascar with 436 firm owners and managers and provides additional evidence that the observed gap in estimates between these two answer formats is caused by a measurement error on the part of survey respondents. Experience in running a business appears to mitigate error but without eliminating it completely.

1 Introduction

Microeconomic research on corruption commonly uses survey data to measure bribery (Reinikka and Svensson, 2006 ; Sequeira, 2012). Despite evidence that respondents to surveys are not always candid when discussing sensitive behaviour such as bribery (Tourangeau and Yan, 2007), only a limited number of studies have intended to estimate the extent to which this lack of candour affects the reliability of bribery indicators constructed from survey data (Azfar and Murrell, 2009; Clausen et al., 2010 ; Kraay and Murrell, 2016 ; Jensen et al. 2010).

Using data from the World Bank's Enterprise Surveys carried out in 15 countries in sub-Saharan Africa between 2006 and 2007[1], Clarke (2011) noticed that a slight change in the answer format to a question about the cost of bribery considerably affects the responses obtained from respondents.

Average estimates by firm managers of the amount requested by public officials to firms similar to theirs in order to make sure that "things get done" lie between 2.5 percent and 4.5 percent of turnover for these 15 countries.[2] With only a few exceptions, Enterprise Surveys implemented before 2005 requested respondents to estimate this amount as a percentage of turnover (Clarke, 2011). Since 2005, surveyees have been offered the options of answering the question either in monetary terms or as a percentage of turnover. Clarke's striking observation is that the average estimates in percentage of turnover are 4 to 15 times higher than estimates in local currency for the 15 sub-Saharan countries investigated. Firm managers who estimated a positive amount of bribes and expressed this amount as a share of turnover reported an average amount of bribes between 4 percent and 8 percent of turnover in most countries. Those who estimated a positive amount of bribes in monetary value evaluated them on average at 0.5 percent to 1.0 percent of turnover.[3]

In his paper, Clarke convincingly dismissed the possibility that observable or unobservable firm characteristics could explain this gap. He also showed that the higher estimates of bribery for firms answering as a percentage of turnover could not result from the influence of extreme values or of more frequent interactions with public officials.

Interestingly, Clarke noted that the response format also appear to influence answers to less sensitive questions. Respondents to Enterprise Surveys carried out in sub-Saharan Africa in 2006 and 2007 consistently reported lower losses due to load-shedding and crime, as well as lower

---

[1] The fifteen countries are Angola, Botswana, Burundi, Democratic Republic of the Congo, Gambia, Ghana, Guinea-Bissau, Guinea-Conakry, Kenya, Mauritania, Namibia, Rwanda, Swaziland and Tanzania.

[2] The question was worded 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 percentage of total annual sales or estimated annual value do establishments like this one pay in informal payments/gifts to public officials for this purpose?"* (The World Bank, 2016a).

[3] Clarke mentioned that this difference is statistically significant at 5 percent for all countries where at least five firms responded to the question as a percentage of turnover and at least five firms responded in monetary value.

security spending when they expressed these amounts in monetary value rather than as a percentage of sales.[4]

Clarke concluded that the only plausible explanation for the difference in the magnitude of bribery estimated in monetary value and as a percentage of turnover is that firm owners and managers tend to overestimate percentages.

Enterprise Survey questions allowing responses in monetary value or as a percentage of turnover, including the one on bribes paid to public officials, explicitly mention both accepted answer formats. Respondents should therefore be in a position to select the answer format they favour.[5] As Clarke (2011) pointed out, however, it is not clear to what extent the choice of the answer format is due to respondent preferences, interviewer preferences, question wording, or questionnaire layout.

Therefore, it may well be that respondents' choice of answer format is a function of their answer. In particular, businesspeople who want to report less bribes could be more inclined to express this amount in monetary value, whereas those wanting to report a larger amount are keen to report it as a percentage of turnover for convenience. This behaviour could explain the observed relationship between the answer format and estimates of bribery.

This paper intends to reveal whether firm owners and managers are truly influenced by the answer format when estimating the magnitude of bribery by randomly assigning to respondents the two competing answer formats. The experiment was designed as part of a business survey implemented by the NGO Transparency International in November 2014.[6] The main objective of the survey was to identify constraints to setting up and running firms operating in Antananarivo, the capital city of Madagascar.

---

[4] Although statistically significant, it should be noted that the difference in estimates between the two answer formats is smaller in size for these other costs compared to bribery.

[5] The answer as a percentage of turnover is usually favoured by respondents to *Enterprise Surveys* when the two response options are available. For all fifteen *Enterprise Surveys* assessed in Clarke (2011), respondents who estimated a positive amount of bribes were over 70 percent to answer the question as a percentage of turnover.

[6] Although more focused on the topic of corruption than the Word Bank's *Enterprise Surveys*, the Transparency International survey questions matching those of the *Enterprise Surveys* were systematically phrased in the same way in order to make the results as comparable as possible. An *Enterprise Survey* was carried out in Madagascar between November 2013 and May 2014, a few months before Transparency International's survey. This *Enterprise Survey* has the same specificity as those carried out in sub-Saharan Africa in 2006 and 2007, i.e. estimates of the annual amount that firms need to pay in gifts and other informal payments to officials to "get things done" are higher when they are reported as a percentage of turnover. Respondents who answered the question in monetary value and provided a positive estimate of bribery evaluated this cost to be 2,080,000 Ariary on average (excluding extreme values distant from more than three standard deviations from the mean value), equivalent to 1.3 percent of their turnover. Those who estimated positive bribes as a percentage of turnover (74 percent of respondents overall) gave an average estimate of 8.6 percent (again outliers excluded). The difference in estimates between the two answer formats is significant at 1 percent.

The rest of the paper is organised as follows: Section II describes the survey, Section III assesses the influence of the answer format on firms' estimates of the magnitude of bribery, Section IV aims at explaining this influence with an empirical model and Section V offers concluding remarks.

2 The data

The survey used for this research interviewed 436 business owners and managers from Madagascar about their experience and perception of corruption in their country (Lesné and Rakotomamonjy, 2015). The first phase of the survey targeted 247 firms created within two years preceding the survey in order to identify the general constraints for setting up a business in Madagascar. The second phase surveyed 189 firms already established for at least two years and with at least ten employees. Both phases of the survey were administered in face-to-face by the same team of interviewers.[7] Most questions were common to the two phases of the survey.

Two versions of the survey questionnaire were randomly allocated to the respondents. The first asked respondents to estimate the average share of the annual turnover of firms similar to theirs that those firms must spend, on average, on gifts and other informal payments to public officials to "get things done".[8] The other version of the questionnaire asked respondents to formulate this amount in monetary terms.[9]

Respondents who had to estimate bribes as a percentage of turnover (monetary value) were also instructed to estimate ten other amounts of interest in percentage (value) throughout the questionnaire.[10]

Two questions asked firm owners and managers about their annual losses due to load-shedding and crime, to be estimated in monetary value or as a percentage of their turnover. Six other questions asked respondents about bribes paid by their firm for power and water hook-ups or to issue building permits, title deeds or import licenses during the two years preceding the survey. These amounts had to be expressed by the respondent either in monetary value or as a percentage of the total cost of those procedures. Another question asked respondents to estimate the monetary value or the average percentage of the cost of litigation that firms like theirs have to spend, on average, to win a case in court. Finally, one question asked firm owners and managers about their perception of the frequency of load-shedding in the 12 months preceding the survey,

---

[7] See Lesné and Rakotomamonjy (2015) for details about the survey methodology.

[8] The exact question was: On average, what percentage of annual turnover do firms like yours pay in informal payments and gifts to public officials to "get things done"?

[9] The exact question was: On average, how much do firms like yours pay annually in informal payments and gifts to public officials to "get things done"?

[10] Interviewers were instructed to ask respondents, if necessary, to rephrase their answer in the format provided by the version of the questionnaire that was allocated to them.

to be answered in number of days or as a percentage of days in the year. The question on bribery was asked toward the end of the survey, after all these questions.[11]

Table 1 shows the characteristics of the respondents and their firms according to the version of the questionnaire administered to them.[12] Descriptive statistics confirm that the observable average characteristics of the two groups of respondent are statistically equivalent.

Table 1: Characteristics of respondents and their firms, per questionnaire version

Proportional share or Average value	Monetary value	Percentage	Difference is null	All sample	Number of observations
Respondent owns the company	51.5%	45.7%	0.235	48.6%	432
Business sector: Retail trade	35.7%	41.3%		38.5%	
Business sector: Manufacturing	13.1%	13.0%	0.464	13.0%	435
Business sector: Services	51.3%	45.8%		48.5%	
Firm has less than 2 years	62.3%	61.8%	0.913	62.1%	436
Number of employees (in logarithm)	1.94	2.00	0.667	1.97	436
Firm's turnover (in billion Ariary)	1.01	0.95	0.885	0.98	297

Tests used for categorical variables are based on Pearson's X^2 statistic with Rao and Scott's (1984) correction. Tests for continuous variables are Adjusted Wald F tests. P-values are reported.

---

[11] Most of these questions were only asked to respondents who had previously admitted at least one recent interaction with the corresponding public service. Only the questions on bribery, load-shedding frequency, and firm losses due load-shedding were asked to all respondents. The response rate to each of these questions, with the exception of those asked indiscriminately to all respondents and the one asking about firm losses due to crime, never exceeds 12 percent of the survey sample. For this reason, only questions that were asked to all respondents and the question about losses resulting from crime are used in the analyses presented in the rest of this paper.

[12] Given the relatively small sample size of the second phase of the survey, Transparency International decided to focus on the version of the questionnaire expressing amounts as a percentage of turnover in order to obtain sufficient statistical accuracy for some of its indicators. As a result, of the questionnaires allocated to the 189 firms who took part in the second phase, 60 percent were of the percentage type and 40 percent were of the monetary value type. This paper takes into account this specificity of the sampling design by systematically including in the analyses a weighting of observations corresponding to the inverse of their likelihood of being drawn. This correction was applied using the Stata *svy* environment and the *pweight* weighting option (StataCorp L.P., 1985). For the first phase of the survey, both versions of the questionnaire were allocated to the 247 respondents on a strict parity basis, and no correction was required.

3. The influence of the answer format on estimates of bribes and other costs

3.1. Average estimates of bribes per answer format

On average, survey respondents estimated that firms similar to theirs must spend 6.2 percent of their turnover in informal payments and gifts to public officials annually, to "get things done".[13] Translated into percentage of turnover, the average estimate of respondents whose version of the questionnaire required an answer in monetary value is 7.7 times lower than the average amount reported by respondents who had to answer the question as a percentage of turnover, a difference significant at 1 percent (Table 2).

These results are in line with Clarke's (2011) observations based on Enterprise Surveys carried out in sub-Saharan Africa in 2006 and 2007. Owners and managers of firms operating in Madagascar provide a much higher estimate of the magnitude of bribery as a percentage of turnover than in monetary value.[14]

As both sub-samples are similar in all respects except for the questionnaire version that was assigned to them, a first conclusion drawn from this experiment is that the magnitude of bribery as measured with survey reports from firm owners and managers is highly sensitive to whether estimates are formulated in monetary terms or as a percentage of turnover.

Table 2: Average estimates of bribes, per questionnaire version

Estimates as a share of turnover	Average estimate	Linearised Standard error	Number of observations
Answers in monetary value	1.19	0.43	56
Answers as a percentage of turnover	9.20	1.41	110
All Answers	6.17	0.95	166

The p-value of the Adjusted Wald test for the difference between estimates of bribery as a percentage of turnover and in monetary value is inferior to 0.001.

---

[13] A similar question is asked in the standard *Enterprise Survey* questionnaire, the only difference being that Transparency International's survey question requires a specific answer format (monetary value or percentage of turnover) whereas the *Enterprise Survey* questionnaire allows for both. As for the *Enterprise Surveys*, extreme values – over or under three standard deviations from the mean – are discarded from the analysis.

[14] The ratio between average estimates formulated as a percentage of turnover and in monetary terms (7.7) is similar to that of the *Enterprise Survey* carried out in Madagascar in 2013 (6.8). It should be noted, however, that this ratio obtained with *Enterprise Surveys* omit null responses as they are not attributable to one or the other of the answer formats. If the analysis of the Transparency International survey data is similarly limited to strictly positive estimates of corruption, the ratio of the average estimates of bribery as a percentage of turnover and in monetary value increases to 9.6.

3.2. The distribution of estimates of bribery per answer format

The distribution of estimates of the magnitude of bribery is very asymmetrical. The proportion of respondents who believe that bribery is totally non-existent is high among those who agreed to answer the question: 25.0 percent for the version of the questionnaire requesting an answer in monetary value and 39.4 percent for the version asking for an answer as a percentage of turnover.[15]

Estimates of bribes expressed in monetary value, once converted as a share of turnover, appear to be heavily concentrated under 1 percent, unlike estimates formulated by respondents in percentage of turnover which are much more disparate (Figure 1). Estimates as a percentage of turnover also reach higher values.[16]

Figure 1: Frequency distribution of bribe estimates, per answer format

Source: Author. Data are from Transparency International
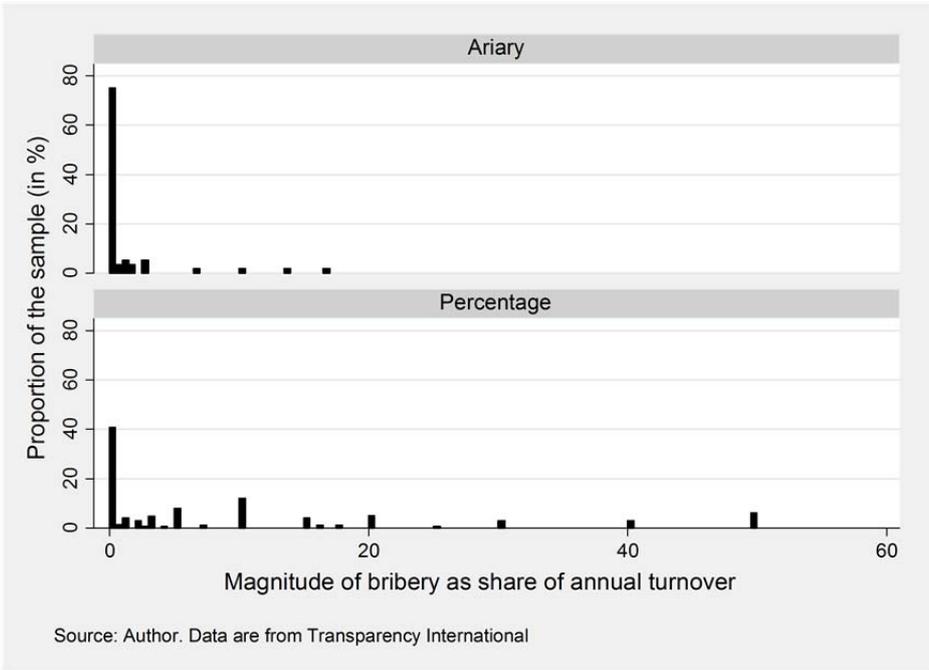
The gap between estimates of bribery expressed in monetary terms and as a percentage of turnover is noticeable for both low and high estimates, as illustrated by the first panel of Figure 2 which displays percentiles of strictly positive estimates of bribery for the two versions of the questionnaire. The ratio between answers as a percentage of turnover and in monetary value for each percentile of the distribution of positive estimates is presented in the second panel of Figure 2. This ratio is systematically higher than one, confirming that the difference in estimates due to the

---

[15] The difference in the proportion of null estimates in the two versions of the questionnaire is statistically significant at 10 percent.

[16] High estimates of bribery as a percentage of turnover is not unique to the Transparency International survey. Among the 15 Enterprise Surveys carried out in Sub-Saharan Africa in 2006 and 2007 mentioned in Clarke (2011), all but one (Guinea Bissau) display a least one estimate of bribery as a percentage of turnover equal or higher than 30 percent.

answer format is not limited to a specific part of the distribution of answers.

Figure 2: Percentiles of bribe estimates, per answer format (first panel) and ratio between estimates as a percentage of turnover and in monetary value, per centile of the distribution of estimates (second panel)

Source: Author. Data are from Transparency International

Source: Author. Data are from Transparency International
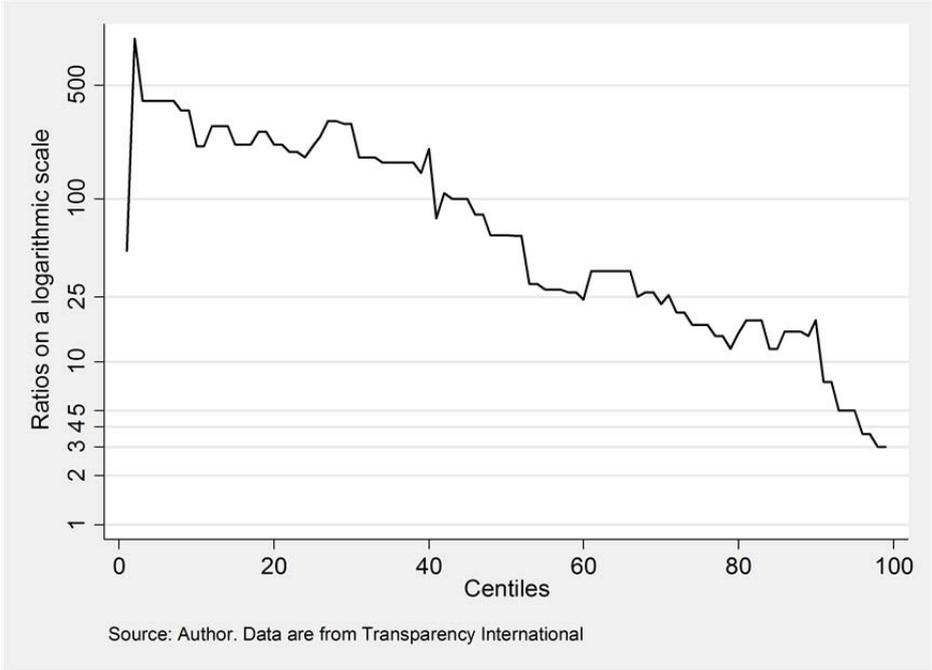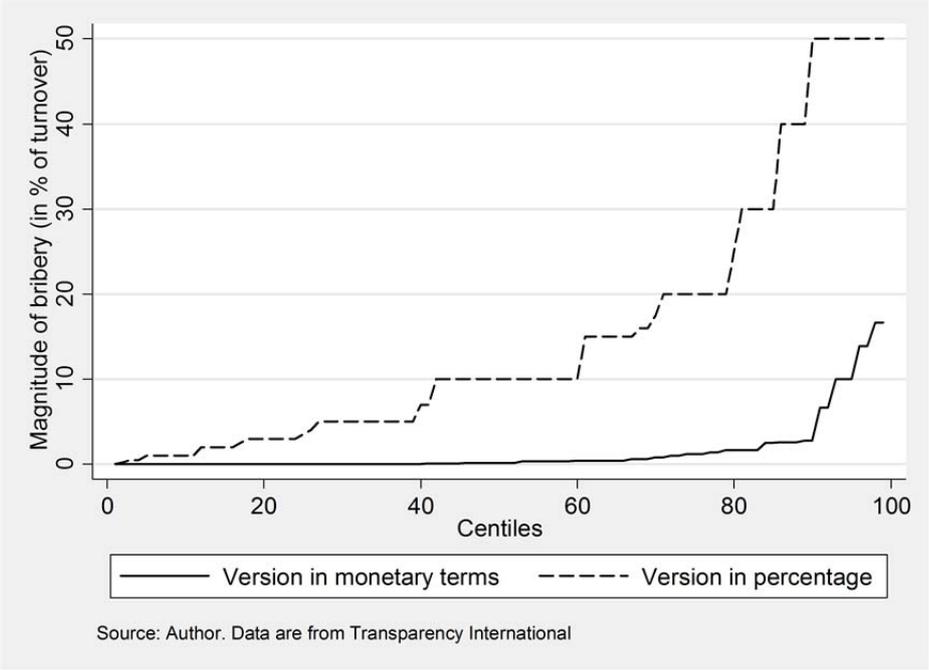
Although the gap is noticeable across the entire distribution of estimates, it appears to be more pronounced for firms reporting a positive but low level of bribes. The ratio of average estimates expressed as a percentage of turnover and in monetary value, excluding zero estimates, decreases from 210 to 18 between the first and the ninth decile of the answer distribution.

3.3 The influence of the answer format on estimates of other costs

The observed gap between the average estimates in monetary terms and as a percentage of turnover is not specific to bribes paid by firms to public officials. It is also apparent for firms' annual losses due to load-shedding and crime.[17] Those losses are of a higher magnitude than bribery: 11.1 percent and 10.1 percent of average firm turnover for load-shedding and crime, respectively, compared to 6.2 percent for bribery. Losses due to crime estimated as a percentage of turnover are 2.7 times higher than when estimated in monetary value.[18] This ratio is 6.0 for losses due to load-shedding,[19] to be compared with a ratio of 7.7 for bribery (Figure3).

The observation that the answer format also affects estimates of less sensitive costs than bribery led Clarke (2011) to conclude that the sensitivity of the topic of corruption - or the phrasing of the question - cannot alone explain the difference in estimates. Clarke nonetheless noted from *Enterprise Surveys* data that despite being statistically significant, the difference in estimates is generally smaller for these other costs. As Figure 3 shows, this is also the case for the Transparency International survey.

More than the sensitivity of the question topic, it may be its abstractness that explains the influence of the answer format on estimates. Losses caused by theft, vandalism, fraud or scams may be less frequent, more striking and therefore more set in respondents' minds than losses caused by power cuts - difficult to measure because they affect many dimensions of the firm's activity - or bribery, which is for many firms of common occurrence.

---

[17] While the question on bribery asked respondents to estimate the amount of bribes paid by firms "similar to theirs", load-shedding and crime questions relate to costs incurred by the firm itself.

[18] Unlike bribery and losses due to load-shedding, the estimation of firm losses due to crime was preceded by a question asking managers if they incurred losses from theft, vandalism, fraud or scams during the 12 months prior to the survey. Approximately half of the managers (49.7 percent) gave a positive answer to this question. They alone had to estimate their firm losses due to crime. The observable characteristics of firm owners and managers who see themselves as victims of criminal acts and of their businesses are similar to the average of respondents, except for a higher proportion of sole proprietorships and businesses operating in the trade sector among the victims.

[19] Estimates of losses due to load-shedding and estimates of losses due to crime expressed as a percentage of turnover and in monetary terms are significantly different at 1 percent.

Figure 3: Average firm costs, per answer format

Source: Author. Data are from Transparency International
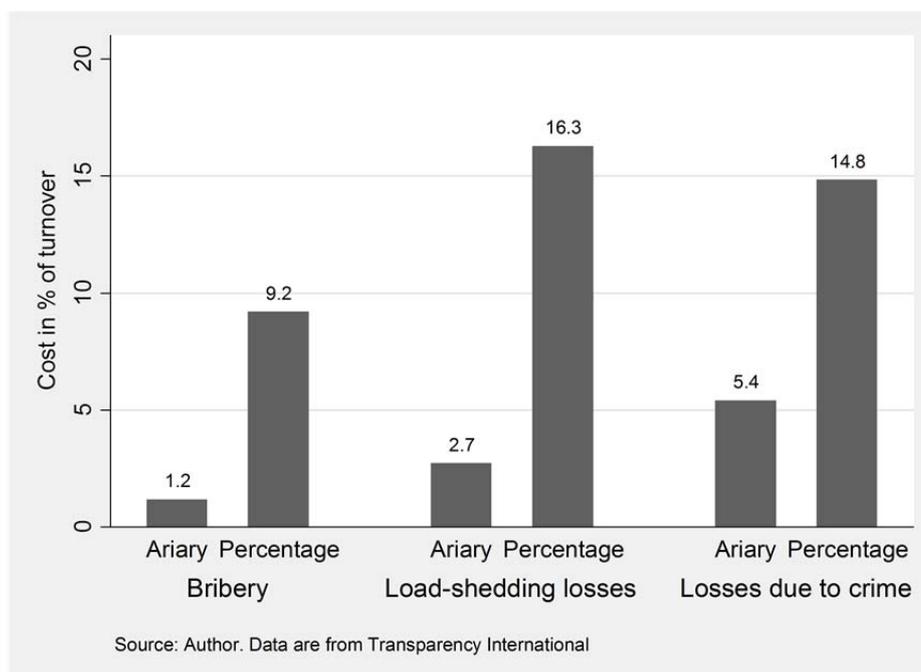
3.4 The influence of the answer format on the absence of response

The question on bribery exhibits a high non-response rate: 47.4 percent of surveyees chose not to answer that question.[20] Non-response, justified by the respondent either as refusal to answer or ignorance of the correct answer, is higher for the questionnaire version asking respondents to estimate bribery in monetary terms (58.8 percent versus 36.0 percent). The difference in non-response rates between the two versions of the questionnaire is significant at 1 percent for this question (Figure 4).

The influence of the answer format on the response rate is even more pronounced for the question about firm losses caused by load-shedding. The non-response rate for the monetary version of the questionnaire is more than twice (57.9 percent versus 26.5 percent) that of the version in percentage. In contrast, the proportion of respondents who did not answer the question on firm losses due to crime in the two versions of the questionnaire is almost identical: 12.7 percent for the version of the questionnaire in monetary terms and 14.1 percent for the questionnaire in percentage. Non-response to this question about crime losses is also much less common than for the other two questions asking about firm losses caused by load-shedding and bribery.

As previously suggested, firm owners and managers may have a more accurate idea of their losses due to criminal acts, compared to how much power cuts and bribery have cost their firm. This may

---

[20] The non-response rate to the same question in the 2013 *Enterprise Survey* carried out in Madagascar is 79.7 percent. Given their sensitive nature, questions assessing perceptions of bribery often exhibit high non-response rates.

account for at least part of the observed difference between the response rates to these questions. For questions about bribery and load-shedding losses, the answer format has a significant influence on the decision to provide an answer. Some firm owners and managers apparently prefer to decline answer those questions in monetary terms, even though they would probably have provided an answer if they had been asked to formulate it as a percentage of turnover.

Evaluating losses due to load-shedding and estimating the magnitude of bribery – two already demanding cognitive processes – might require even more thought in monetary value than as a percentage of turnover. This additional reasoning effort could discourage some surveyees from answering these questions. Consistent with this, Iarossi (2006) proposed that "[i]t might be easier for respondents to provide answers in categories or percentages rather than in absolute values". Preference for this answer format is confirmed by the slightly higher level of cooperation experienced by interviewers from surveyees who responded to the version of the questionnaire in percentage.[21]

Figure 4: Non-response rates to questions about costs, per answer format

Source: Author. Data are from Transparency International

---
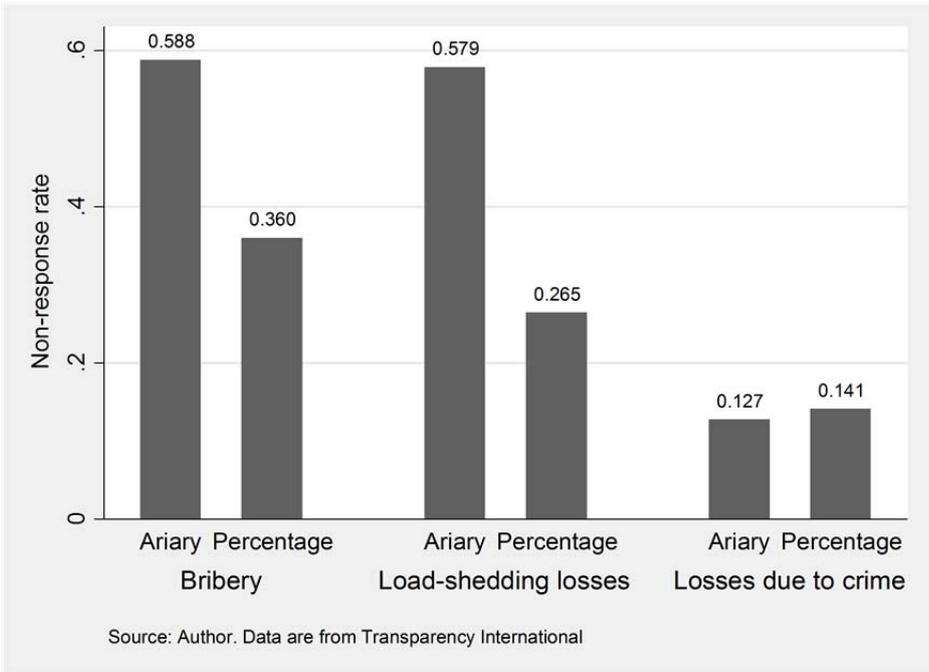
[21] The level of cooperation, defined as the respondent's willingness to answer sincerely all the survey questions, was noted by the surveyors at the end of their interviews on a scale from 0 (no cooperation) to 10 (full cooperation). The average cooperation score for surveyees who had to answer the questionnaire in monetary terms was 7.3 and 7.7 for those who answered the questionnaire in percentage. The difference is significant at 10 percent.

4. An empirical model

4.1. What we have learned so far

On average, firm owners and managers estimated bribes paid to public agents by firms similar to theirs to be 7.7 times higher when they formulated their answer as a percentage of turnover rather than in monetary value.

A significant difference is observed throughout the distribution of positive estimates, but in a magnitude that decreases as estimates of bribery rise. Also, the difference in estimates as a percentage of turnover and in monetary value is not specific to the question asking survey respondents to estimate how much firms similar to theirs must pay in bribes to "get things done". Estimates as a percentage of turnover are also significantly higher than estimates in monetary value for firm losses due to load-shedding and crime. The more abstract the measured concept is, the larger is the gap between estimates as a percentage of turnover and in monetary value.

Respondents generally favour the answer format in percentage. The non-response rate to the bribery and load-shedding questions is significantly lower for this version of the questionnaire.

4.2. Where may the difference in estimates come from?

The influence of the answer format on estimates of the magnitude of bribery may be due to respondents overestimating amounts expressed as a percentage of turnover, underestimating them in monetary value, or a combination of both types of error.[22] The first type of error may originate from the fact that some business owners and managers are not comfortable in manipulating percentages, as Clarke (2011) suggested. The second type of error may result from the fact that some survey respondents answering in monetary value are not exhaustive in their estimations, merely focusing on a few cases where bribes or other costs were particularly noticeable, leaving out information that may be relatively harder to gather for the respondent, although still relevant for the overall estimation requested by the question.

Firm owners and managers with less experience in running a business may be more subject to making errors when estimating the magnitude of bribery in monetary value or as a percentage of turnover, being perhaps less used to juggling figures. The first phase of the Transparency International survey specifically targets newly-created firms, providing a useful framework to test this hypothesis.

---

[22] Interviewers rated similarly their perceived reliability of estimates produced by respondents in percentage and in monetary value. Using a calculator was no more frequent among respondents of one version of the questionnaire than the other. Paper and pen were use more often by respondents who were allocated the questionnaire in monetary terms (5.5 percent versus 2.3 percent, significant at 10 percent). Conversely, a larger share of respondents in percentage did consult their records compared to respondents in monetary terms (19.2 percent versus 13.1 percent, significant at 10 percent).

4.3 The model

Using Transparency International's survey, I examine whether experience in running a business influences the relationship between the answer format and estimates by business owners and managers of the amount of bribes paid annually by firms similar to theirs.

Since a significant share of respondents estimated that firms do not need to pay any bribes, a two-part model for mixed discrete-continuous outcomes is appropriate. Two-part models are limited dependent variable models exhibiting a non-trivial mass of observations reaching a value that is the lower bound those observations can possibly assume.

In my two-part model, the probability of observing a strictly positive outcome is modelled using a binary choice model. The second part of the model fits a regression model on the continuous range of positive outcomes. This two-part model is a flexible alternative to the Tobit model in some instances when the latter is unsuitable (Cragg, 1971). This is the case when a key assumption of the Tobit model does not hold. In a Tobit model, the same mechanism must explain both the binary outcome (positive versus zero) and the magnitude of the effect on positive outcomes. Consequently, any explanatory variable that affects positively the probability of obtaining a non-zero outcome must also have a positive effect throughout the range of non-zero outcomes (Wooldridge, 2010).

As shown in the previous section, the proportion of zero outcomes is higher for the version of the questionnaire asking surveyees to estimate bribes as a percentage of turnover. Conditional of the estimation being strictly positive, estimates of the magnitude of bribery as a percentage of turnover are higher, on average, than estimates in monetary terms. Therefore, in this case, the common mechanism hypothesis of the Tobit model is unlikely to hold. Two-part models, also called hurdle models in the case of count variables (Cameron and Trivedi, 2013), allow for different processes to explain the probability of a positive versus zero outcome and to model the function explaining strictly positive outcomes.

The two-part model used in this paper combines a probit model with a linear regression model on strictly positive estimates of bribery. Both parts are modelled using as explanatory variables the firm's sector (retail trade, manufacturing, services) and the logarithm of the number of employees. The same explanatory variables were used in Clarke (2011) as a base model to control for firms' observable characteristics when assessing differences in estimates of bribery in monetary value and as a percentage of turnover on Enterprise Surveys data.

Columns 1 and 2 of Table 3 show the coefficients of the probit regression of the zero versus positive estimates of bribery and the OLS regression on positive estimates as a function of the base model explanatory variables. Column 3 reports average marginal effects. Retail trade firms estimate bribes as a percentage of turnover that are 4.4 percentage points higher, on average, than firms operating in the services and retail trade sectors (significant at 10 percent). The size of the firm, measured by the logarithm of the number of employees, is not a significant predictor of bribes.

Unsurprisingly, asking respondents to estimate bribes as a percentage of turnover rather that in monetary value increases significantly their average estimation. Columns 4 to 6 of Table 3 display the coefficients of the probit and OLS regressions as well as average marginal effects for all explanatory variables from the base model to which is added a dummy variable indicating the version of the questionnaire assigned to the respondent (augmented model). Answering as a percentage of turnover increases estimates of the amount of bribes paid by 7.7 percentage points, all other things held equal. The null hypothesis that the average marginal effect of the answering format on estimates of bribery is 0 is rejected at 1 percent.[23]

Table 3: Empirical results for the base and augmented models

VARIABLES	(1) Base Model Probit Coefficients	(2) Base Model OLS Coefficients	(3) Base Model Average marg. effects	(4) Augmented Model Probit Coefficients	(5) Augmented Model OLS Coefficients	(6) Augmented Model Average marg. effects
questionnaire (in %)				-0.451**	14.23***	7.682***
				(0.228)	(2.092)	(1.635)
log(employees)	0.119*	0.0633	0.444	0.141*	-0.782	0.0202
	(0.0712)	(0.867)	(0.610)	(0.0727)	(0.732)	(0.561)
sector: services	-0.258	6.060**	3.298	-0.246	5.224**	2.714
	(0.333)	(2.952)	(2.109)	(0.335)	(2.631)	(1.999)
sector: retail trade	0.0396	6.037*	4.367*	0.0172	6.660**	4.741*
	(0.342)	(3.469)	(2.566)	(0.344)	(3.190)	(2.470)
Constant	0.310	3.971		0.563	-2.340	
	(0.354)	(3.470)		(0.378)	(3.158)	
R^2		0.0225			0.2823	
McFadden R^2	0.0225			0.0430		
Observations	166	166	166	166	166	166

Standard errors in parentheses;
*** p<0.01, ** p<0.05, * p<0.1
The two-part model was estimated with Stata (Belotti et al., 2015)

Adding a binary variable coding 1 if the firm has been created less than two years ago, and 0 if the firm has two years or more gives us the full model presented in columns 1 to 3 of Table 4. Being the owner or the manager of a newly created firm, a proxy indicator for the respondent's experience in running a firm, does not affect the likelihood of respondents admitting the existence of bribery. It does, however, significantly increase the estimation of the magnitude of bribery for those respondents who estimated a positive amount of bribes. The average marginal effect for this variable is large, comparable in size to that of the questionnaire version: respondents owning or managing a new firm estimate bribes to be 7.6 percentage point higher, *ceteris paribus*, than respondents whose firm had been established at least two years before the survey was carried out.

---

[23] In addition, the augmented model confirms the unsuitability of the Tobit model as an estimation model in this present case since the version of the questionnaire has an opposite effect on the propensity to admit bribery and on the estimation of bribes conditional on the respondent admitting the existence of bribes.

If respondents' experience in running a business is correctly measured by their firm being at least two years old, this result could mean that, all other things held equal, respondents with less experience estimate more bribes. Alternatively, this result could simply come from the fact that new firms pay more bribes than older firms.

To decide between those two differing interpretations, an interaction term between the version of the questionnaire and the "new firm" binary variable is added to the full model. Since assessing the statistical significance of an interaction term in a non-linear model is not straightforward (Ai and Norton, 2003; Greene 2010), I focus on the second step of the two-step model, i.e. on the outcomes of the linear regression run on strictly positive estimates of bribery.[24] As Table 4 shows, the interaction term is large in magnitude and statistically significant. Among respondents admitting that firms like theirs must pay bribes, owners and managers of newly created firms who answered as a percentage of turnover provided estimates of bribery that are 11.8 percentage points higher than those who answered the question in monetary terms, *ceteris paribus*. This is almost twice the average estimation of the magnitude of bribery in the sample (6.2 percent).

With the interaction term, the "new firm" binary variable is no longer significant on its own, suggesting that the positive effect of that variable in the full model was not due to the fact that the cost of bribery is larger for younger firms than for older ones. The questionnaire version variable is still highly significant even after the inclusion of the interaction term, but the magnitude of the average marginal effect has been reduced in half, from 13.4 to 6.3 percentage points. Although the business experience of respondents as measured by the two-year cut-off for their firm's age appear to explain part of the gap between estimations of bribery as a percentage of turnover and in monetary terms, the answer format is still strongly associated with the magnitude of estimates even for older firms.

---

[24] The proportion of strictly positive estimates of bribery among surveyees who did not decline to answer the question is superior for the questionnaire in monetary terms for both younger and older firms, in a magnitude that is comparable (79.6 percent versus 70.9 percent for new firms, and 82.9 percent versus 71.5 percent for older firms). This result indicates that the interaction term is unlikely to play a significant role in the first part of the two-part model.

Table 4: Empirical results for the full model

VARIABLES	(1) Full Model Probit Coefficients	(2) Full Model OLS Coefficients	(3) Full Model Marginal effects	(4) Full Model with interaction Probit Coefficients	(5) Full Model with interaction OLS Coefficients
questionnaire (in %)	-0.450*	13.37***	7.125***	-0.528	6.276***
	(0.228)	(1.972)	(1.566)	(0.392)	(2.085)
log(employees)	0.131	2.296	2.018*	0.131	2.050
	(0.120)	(1.506)	(1.144)	(0.120)	(1.374)
sector: services	-0.241	3.777	1.687	-0.246	3.786*
	(0.338)	(2.557)	(2.017)	(0.336)	(2.285)
sector: retail trade	0.0236	4.066	2.935	0.0200	4.656*
	(0.353)	(2.855)	(2.273)	(0.351)	(2.555)
newfirm	-0.0386	11.76**	7.623**	-0.116	4.175
	(0.367)	(4.716)	(3.588)	(0.521)	(3.711)
newfirm*questionnaire				0.112	11.78***
				(0.479)	(3.513)
Constant	0.600	-13.43**		0.661	-8.413*
	(0.511)	(5.732)		(0.567)	(4.871)
R^2		0.3425			0.3856
McFadden R^2	0.0430			0.0433	
Observations	166	166	166	166	166

The gap in estimates of bribery as a percentage of turnover and in monetary terms is wider for new firms. This is also the case for the estimation of firm losses due to crime, as shown by Figure 5. The ratio between estimates of crime costs expressed as a percentage of turnover and in monetary value is 3.1 for new firms and 1.9 for order firms[25]. However, for firm losses caused by load-shedding, the ratio of average estimates as a percentage of turnover and in monetary terms is not larger for new firms (5.9) than for old firms (6.3).

---

[25] Interestingly, an adjusted Wald test does not reject the null hypothesis of equality of average estimates for crime losses as a percentage of turnover and in monetary value when the sample is limited to firms already created two years prior to the survey.

Figure 5: Cost estimates by new and older firms, per answer format

Source: Author. Data are from Transparency International
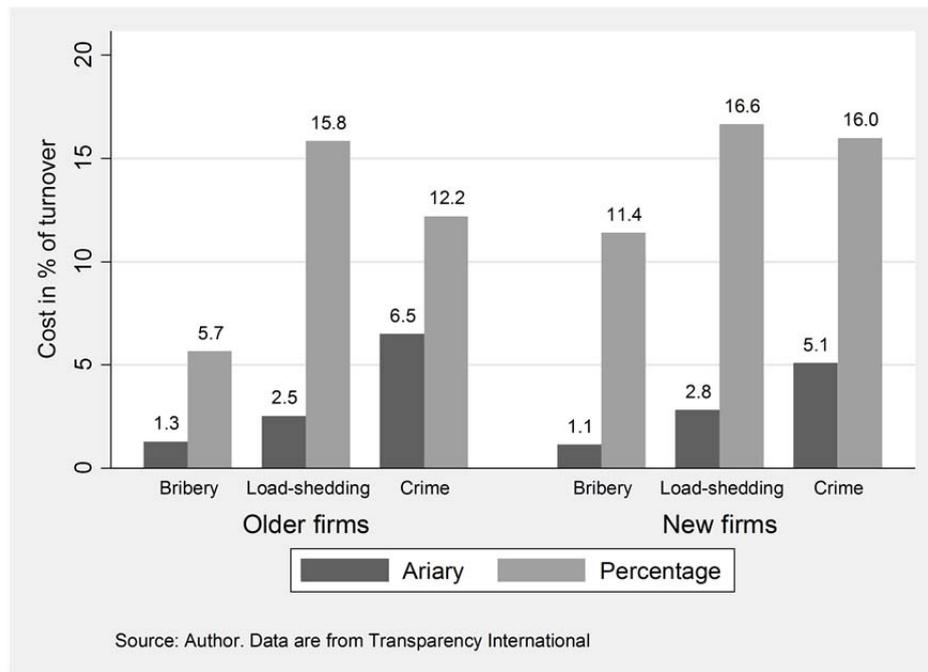
4.4. Selection

Following Clarke (2011), to enable comparisons between the estimates of magnitudes of bribery and other costs from the two versions of the questionnaire, I converted estimates expressed in monetary value as a percentage of turnover using the firm's own annual turnover as reported by the firm owner or manager during the survey.[26]

A concern is that some respondents may under-report their firm's turnover, particularly if they fear that the information they provide would be shared with the tax authorities. If so, their estimates of bribery converted into percentage of turnover would be biased downwards. This under-reporting of turnover would have the effect of reducing the observed gap between the average estimates of bribery and other costs expressed in monetary value and as a percentage of turnover. Without this bias, the gap between estimates in monetary terms and as a percentage of turnover could possibly be wider than the one observed.

A more problematic issue raised by the conversion is that some respondents may estimate bribes for firms that are systematically smaller in size, despite the question asking them to evaluate how much *firms similar to theirs* have to pay annually in bribes to "get things done". This bias is unlikely to explain the observed gap between estimates of bribery as a percentage of turnover and in

---

[26] For firms created less than two years before the survey, the survey asked respondents to estimate their average monthly turnover instead of annual turnover.

monetary terms, as a similar gap – although smaller in magnitude – is noticeable for estimations of crime and load-shedding costs borne by the firm. Those costs, when estimated as a percentage of turnover, relate not to other firms' turnover as for the bribery question but to the firm's own turnover.

The conversion was not possible for estimates of bribery and other costs when respondents in monetary terms who estimated a positive amount were unable or unwilling to reveal their firm's turnover to the interviewer. To maintain comparability, sample observations for which no reporting of the firm's turnover is available (139 observations) have been dropped from all regressions and computations of the magnitude of bribery and other costs.[27]

Sample selection issues may occur if availability of data about turnover is non-random. Also, while the experiment has randomly allocated the two versions of the questionnaire (monetary value versus percentage of turnover) to two sub-sample of surveyees with identical average characteristics, some respondents have declined responding to the question. As noted in the previous section, the non-response rate to the bribery question is significantly lower for the version of the questionnaire in percentage. The same difference in the non-response rate of the two versions of the questionnaire has been noticed for the question about the cost of load-shedding, whereas the non- response rate to the question about the cost of crime borne by the respondent's firm is not significantly different for one version of the questionnaire compared to the other.

Interestingly, the gap in estimates in monetary value and as a percentage of turnover is relatively smaller for the question asking respondents to estimate their firms' losses due to crime. Might it be then that firm owners and managers who failed to answer the bribery question in monetary terms would have estimated a lower-than-average estimate of bribery as a percentage of turnover, thus explaining the difference observed between the two answer formats?[28]

To evaluate the influence of selection on the observed difference in the magnitude of bribery estimated in monetary value and as a percentage of turnover, I apply a Heckman correction to the full version of the two - part model including the interaction term between the questionnaire version variable and the "new firm" binary variable. Selection may arise because some respondents refrained from answering the bribery question or because they did not want to or could not

---

[27] A less conservative approach is to leave out only those observations for which the conversion was necessary but not possible due to the respondents failing to report their firm's turnover. For respondents who answered in percentage, their firm's turnover is not of use in the analyses. The firm's turnover is not required either when respondents in monetary terms reported a zero magnitude or did not provide an estimation. Results remain largely unchanged when only the 22 observations for which the conversion was necessary but not possible are dropped instead of the 139 observations for which no turnover data are available. The only noticeable difference is that the ratio between average estimates of firm losses due to load-shedding as a percentage of turnover and in monetary value becomes larger for new firms (7.3) than for older firms (6.4).

[28] This hypothesis is consistent with the fact that the ratio between estimates as a percentage of turnover and in monetary value is slightly lower for the *Enterprise Survey* carried out in Madagascar, as business leaders who took part in this survey may have had the choice of the answer format for this question.

provide information about their firm's turnover.

A probit model with sample selection (Van de Ven and Van Praag, 1981) is performed on the first part of the two-part model and estimated with a maximum-likelihood estimator.[29] A binary variable capturing whether the survey respondent is the owner of the firm is used as an exclusion restriction. The logic behind adding this variable in the selection equation is that some respondents who are not owners but employees of their firm could be unwilling to reveal sensitive information such as the firm's turnover or how much bribes need to be paid without knowing whether the owner of their firm would agree with this disclosure. Respondents who are owners of their firm are significantly (at 5 percent) more likely to provide information on their turnover and to make an estimation of the amount of bribes firms similar to theirs have to pay annually to public officials, on average, to "get things done" (Table 5).

The correlation coefficient between the error terms of the selection and outcome regressions (ρ) is not significantly different from 0 at conventional levels,[30] which is evidence that selection is not a threat to the results obtained on the probit part of the two-part model.

The Heckman selection model applied to the OLS part of the two- part model and estimated with maximum likelihood, again with the owner variable as the exclusion restriction, fails to converge. Moreover, the two- step version of the Heckman selection model is not suitable for a complex survey dataset (see footnote 12). A solution to this problem is to randomly drop 40 surveyees from the large firm component of the survey to which the percentage version of the survey questionnaire had been allocated in order to redress comparability between the two versions of the questionnaire. The two-step approach can then safely be applied to this modified dataset.

In order to use all available information efficiently, this process is repeated multiple times with a new random sub-sample of 40 observations being dropped each time and added back again to the sample after the two- step Heckman selection model has been estimated. With 100,000 iterations, the average p-value obtained for the coefficient of the inverse Mills ratio is 0.71, indicating that selection is also unlikely to be an confounding factor for statistical results obtained with the second part of the two-part model.

---

[29] The *heckprob* command in Stata.

[30] For numerical stability during the optimization process and to ensure that the correlation coefficient ρ is confined within the [-1,1] range, Stata includes atanh(ρ), the inverse hyperbolic tangent of ρ, instead of ρ when estimating a probit model with sample selection. Significance tests over ρ must be applied to atanh(ρ). Since atanh(0) = 0, the reported test for atanh (ρ) = 0 is equivalent to the test for ρ = 0 (StataCorp, 2011).

Table 5: Coefficients of the probit part of the two-part model with sample selection

VARIABLES	(1) Outcome equation	(2) Selection equation
questionnaire (in %)	-0.737**	0.631***
	(0.321)	(0.200)
log(employees)	0.127	-0.0196
	(0.107)	(0.0778)
sector: services	-0.223	0.0544
	(0.296)	(0.203)
sector: retail trade	-0.0211	0.0766
	(0.309)	(0.212)
newfirm	-0.138	-0.00975
	(0.464)	(0.274
newfirm*questionnaire	0.203	-0.189
	(0.396)	(0.259)
owner		0.312**
		(0.139)
Constant	1.393**	-0.748**
	(0.555)	(0.365)
Atanh(rho)	-0.974	
	(1.492)	
Observations	431	431

4.5 An additional test: load-shedding frequency

If estimates of bribery vary depending on the answer format due to a measurement error, this error should affect any type of percentage versus absolute value calculations, not just those involving firm turnover.

One of the questions in the Transparency International survey asked respondents to estimate the number of load-shedding days over the 12 months period preceding the survey.[31] Those who had to estimate the magnitude of bribery as well as business losses due to crime and load-shedding as a percentage of turnover had to answer this question as a percentage of working days in the year. Concretely, this calculation requires dividing the number of load-shedding days experienced during the year by 3.13.[32]

Firm owners and managers estimated an average of 76.5 days of load-shedding affecting their business over the 12 months period preceding the survey, or 24.4 percent of the total number of working days in the year. Those responding as a percentage of days produced a higher estimate

---

[31] Load-shedding – or *délestage* in French-, the planned shut-down of the power supply due to consumption exceeding the production capacity, is a major problem in Madagascar. The World Bank's 2016 Doing Business report attributed the lowest possible score - 0 out of 8 points - for reliability of the power supply and transparency of tariffs from the perspective of businesses operating in Madagascar (The World Bank, 2016b).

[32] This figure is obtained by subtracting from the 365 days in the year 52 days corresponding to non-working days (Sundays), and dividing by 100.

(28.5 percent or 89.3 days) than those responding in number of days (20.2 percent or 63.3 days). The adjusted Wald test rejects at 1 percent the null hypothesis of equality of the mean estimates of load-shedding frequency as a percentage of days and in number of days.

As for the magnitude of bribery and firm losses due to load-shedding, the non-response rate is higher for respondents in number of days, 9.4 percent versus 3.2 percent for those answering as a percentage of days (a difference significant at 5 percent). The proportion of zero estimates is statistically equivalent: 1.4 percent for respondents in number of days and 0.9 percent for respondents as a percentage of days.

The ratio of the estimated frequency of load-shedding expressed as a percentage of days and in number of days is significant, although of a lower magnitude (1.4) than for the bribery, crime and load-shedding costs. This result is consistent with the hypothesis of a computational error causing the difference in estimates.

Less experienced firm owners and managers, who started their business in the 24 months prior to the survey, estimated load-shedding frequency about twice as high in percentage of days as in number of days (Figure 6), a difference statistically significant at 1 percent. For firms at least two years old, the answer format does not influence the average estimate of load-shedding frequency. This frequency is estimated to be 25.0 percent of the number of working days in the year for those who answered in number of days, compared with 26.3 percent for estimates in percentage of days.[33]

Even for this rather simple calculation, the answer format significantly influences estimates of less experienced owners and managers. This result suggests that experience makes businesspeople more competent to formulate such estimates. Experience is not, however, a perfect cure for miscalculations. As shown in the previous section, the difference in estimates of the magnitude of bribery, firm losses due to load-shedding and firm losses due to load crime in monetary value and as a percentage of turnover are significant both statistically and economically even for firms older than two years.

The survey question about load-shedding frequency provides additional insight on the comparative reliability of each answer format. Actual data on the number of load-shedding days recorded in Antananarivo during the year 2014 could be obtained from Jiro sy Rano Malagasy (JIRAMA)[34], the only company supplying power in Antananarivo. As it was not possible to obtain the same data for the months of November and December 2013, the year 2014 is used as a proxy for the frequency of load-shedding during the twelve months prior to the survey carried out by

---

[33] The p-value of the adjusted Wald test for the difference in estimates is 0.77.

[34] I thank JIRAMA for those data.

Transparency International between the 3rd and 28th of November, 2014.[35]

From JIRAMA records, the number of working days on which at least one scheduled power cut occurred in Antananarivo was 74, or 23.6 percent of working days. This figure is slightly lower than the average estimate by firm owners and managers of 76.5 load-shedding days affecting their business over the 12 month period preceding the survey. The real number of load-shedding days is well below the average estimate in percentage (89.3 days), and significantly higher than the average estimate in number of days (63.3 days).

It is worth noting that load-shedding on a given day does not necessarily impact all neighbourhoods throughout the city. For this reason, it is likely that a number of surveyees who estimated a number of days below 74 or 23.6 percent of the total number of working days nonetheless overestimated the number of load-shedding days affecting their business. This is consistent with the fact that firms with at least 10 employees provide an average estimate of the number of load-shedding days affecting their business that is higher than the average estimation obtained from newly created firms, which are generally of a smaller size and operating from a single workplace (80.5 versus 73.7 days, respectively).

Figure 6: Estimation of the proportion of load-shedding days in the year by new and older firms, per answer format

Source: Author. Data are from Transparency International
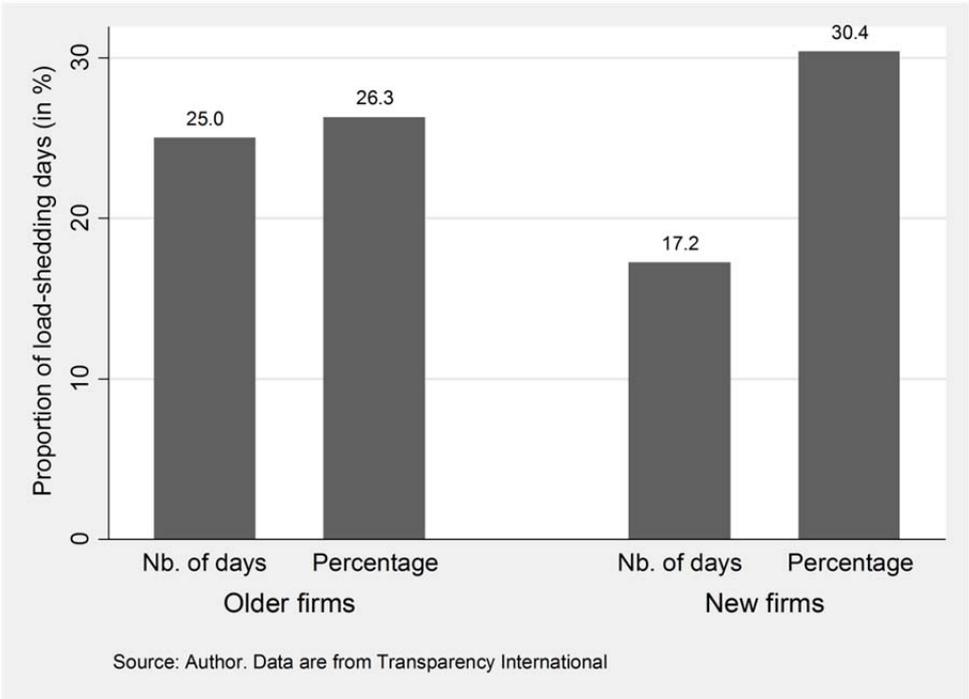
An informative statistic that can easily be computed is the proportion of respondents who

___

[35] According to informed JIRAMA technicians I interviewed, they is no evidence of a different frequency of load-shedding in the months of November and December 2013 compared to November and December 2014.

overestimated with certainty the number of load- shedding days in the year, i.e. the proportion of firm owners and managers who estimated a number of load-shedding days higher than 74 or 23.6 percent of the total number of working days.

Approximately one-fourth (28.2 percent) of respondents who estimated load-shedding frequency in number of days produced an assessment higher than the actual number of load-shedding days in Antananarivo city. In comparison, about half (51.4 percent) of firm owners and managers who answered the question as a percentage of the number of working days in the year overestimated with certainty the number of load-shedding days affecting their firm. This result confirms Clarke's (2011) insight that some survey respondents tend to systematically overestimate percentages. This evidence, however, does not exclude the possibility that respondents to business surveys may also tend to underestimate amounts when formulating them in absolute value.

5. Conclusion

Why do business owners and managers, especially the ones with less experience in running a business, are so much misjudging the magnitude of bribery? A possible explanation is that some respondents to business surveys tend to systematically overestimate amounts in percentages, as suggested by Clarke (2011). An alternative hypothesis is that certain respondents answering in monetary value are not exhaustive in estimating bribes, focusing only on a few payments so as not to have to think too long about their answer, or for lack of interest in producing a reliable answer. Experience in running a business appears to mitigate this inclination to error but without eliminating it entirely.

This research paper is unable to determine which of the two answer formats tested produces the most reliable estimates, only suggesting that some respondents – particularly the less experienced ones – most likely overestimate percentages. Providing a definite answer to this question would require a better understanding of the cognitive mechanisms involved when survey respondents estimate amounts. In this respect, the non-response rate certainly has an important role that needs to be better understood, especially to determine whether there is a trade-off between a higher response rate and more reliable answers.

However, this research is able to confirm the existence of a systematic bias in estimates by firm owners and managers of the magnitude of bribery, and also of other quantities of interest such as firm losses due to load-shedding and crime, and even the frequency of load-shedding. It appears that the difference between estimates in percentage and in absolute value is particularly high for abstract concepts that are difficult to quantify. For bribery more specifically, the difference in estimates as a percentage of turnover and in monetary terms is observed for high and low estimates alike, although the difference appears more pronounced in relative terms for positive but small estimates.

The study of the influence of the answer format on estimations of bribery in business surveys has

clear implications for research on the determinants and effects of corruption. Results of microeconomic studies using business survey data may be significantly affected by measurement error related to the format in which the survey questions are answered. As this research paper shows, such measurement error may affect businesspeople differently depending on their characteristics such as their experience in running a business. Studies using business survey data finding a negative correlation between the longevity of firms and the share of firm turnover paid in bribes may face invalid conclusions if no proper attention is given to the influence of the response format on estimates of bribes.

References

- Ai, C. and Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics letters*, 80(1):123–129.

- 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.

- Belotti, F., Deb, P., Manning, W. G., Norton, E. C., et al. (2015). Two pm: Two-part models. *Stata Journal*, 15(1):3–20.

- Cameron, A. C. and Trivedi, P. K. (2013). *Regression analysis of count data*, volume 53. Cambridge university press.

- Clarke, G. R. (2011). How petty is petty corruption? Evidence from firm surveys in Africa. *World Development*, 39(7):1122–1132.

- 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.

- Cragg, J. G. (1971). Some statistical models for limited dependent variables with application to the demand for durable goods. Econometrica: *Journal of the Econometric Society*, pages 829–844.

- Greene, W. (2010). Testing hypotheses about interaction terms in nonlinear models. *Economics Letters*, 107(2):291–296.

- Iarossi, G. (2006). The power of survey design: A user's guide for managing surveys, interpreting results, and influencing respondents. World Bank Publications.

- 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.

- Lesné, F. and Rakotomamonjy, J. (2015). *Corruption dans les services publics d'Antananarivo : résultats des enquêtes d'entreprise*. Transparency International Madagascar.

- Rao, J. N. and Scott, A. J. (1984). On chi-squared tests for multiway contingency tables with cell proportions estimated from survey data. *The Annals of Statistics*, pages 46–60.

- Reinikka, R. and Svensson, J. (2006). Using micro-surveys to measure and explain corruption. *World Development*, 34(2):359–370.

- Sequeira, S. (2012). Advances in Measuring

Corruption in the Field. Chapter 6 in *New advances in experimental research on corruption*, pages 145–175. Emerald Group Publishing Limited.

- StataCorp (1985). *Stata survey data reference manual*. College Station, TX: Stata Press.

- StataCorp (2011). *Stata 12 Base Reference Manual*. College Station, TX: Stata Press.

- The World Bank (2016a). Enterprise surveys: Indicator descriptions.

- The World Bank (2016b). Doing Business: Measuring regulatory quality and efficiency.

- Tourangeau, R. and Yan, T. (2007). Sensitive questions in surveys. *Psychological bulletin*, 133(5):859.

- Van de Ven, W. P. and Van Praag, B. M. (1981). The demand for deductibles in private health insurance: A probit model with sample selection. *Journal of econometrics*, 17(2):229–252.

- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

"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

FeRDi

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