CMI WORKING PAPER NUMBER 5
REPOA WORKING PAPER 18/001

WORKING PAPER

MARCH 2018

This REPOA/CMI working paper features research on Tanzania as a future petro state

AUTHORS
Alexander W. Cappelen
Odd-Helge Fjeldstad
Donald Mmari
Ingrid Hoem Sjursen
Bertil Tungodden

Understanding the resource curse: A large-scale experiment on corruption in Tanzania

Understanding the resource curse: A large-scale experiment on corruption in Tanzania

CMI Working Paper number 5 2018
Repoa number 18/001
March 2018

Authors

Alexander W. Cappelen
Odd-Helge Fjeldstad
Donald Mmari
Ingrid Hoem Sjursen
Bertil Tungodden

Cover photo

Alessio MumboJumbo on Flickr, CC license.

ISSN 0804-3639 (print)
ISSN 1890-5048 (PDF)
ISBN 978-82-8062-687-5 (print)
ISBN 978-82-8062-688-2 (PDF)

www.cmi.no

Understanding the resource curse: A large-scale experiment on corruption in Tanzania*

Alexander W. Cappelen, Odd-Helge Fjeldstad, Donald Mmari, Ingrid Hoem Sjursen and Bertil Tungodden†

Abstract

Corruption appears to be an important driver of the resource curse in developing countries. We report from a large-scale field experiment in Tanzania that provides causal evidence on how expectations about future natural resource revenues shape expectations about corruption and the willingness to engage in corrupt behavior. We find robust evidence that information about the discovery of natural gas in Tanzania causes people to expect more corruption in the future, but has no impact on people's willingness to engage in corrupt behavior. We believe that our results shed some light on underlying mechanisms of the resource curse.

Keywords: Natural resources, Corruption, Field experiment, Behavioral economics, Political economy, Tanzania

JEL Classification: Q32, Q33, D90, D73, P16

*We would like to thank Assistant Researcher Cornel Jahari for the impeccable organization of the data collection and valuable inputs; Miriam Mganga, Maryam Mnkande, Berrt K. Hurbert, Betty Mjeme, Philbert Karia, Joyce Mamuya, Tumsifu Heriel, Joseph Lucas, Goodluck Nyanda, Albina Kaunda, Segere Mtundi and Sammy Lema for research assistance; Jørgen Juel Andersen, Kendra Dupuy, Ivar Kolstad, Halvor Mehlum, Kalle Moene, Ragnar Torvik, Tina Søreide, and Gerhard Toews and other staff at Oxford Centre for the Analysis of Resource Rich Economies (OxCarre); James Cust, Oliver Morrissey and various seminar and conference participants for valuable inputs at different stages of the project; and administrative staff at REPOA and the Department of Economics for support. We would also like to thank Anton Fouquet and Seme Nassei for the design of the informational videos. The research was funded by the Royal Norwegian Embassy in Dar es Salaam through the research programme "Tanzania as a future petrostate" and was also partially supported by the Research Council of Norway through its Centres of Excellence Scheme, FAIR project No 262675.

†Affiliations: Cappelen & Tungodden: FAIR - The Choice Lab, Norwegian School of Economics, Bergen, Norway; Fjeldstad: Chr. Michelsen Institute, Bergen, Norway, and African Tax Institute, University of Pretoria, South Africa; Mmari: REPOA, Dar es Salaam, Tanzania; Sjursen: Centre for Applied Research at NHH (SNF), Norwegian Centre for Taxation (NoCeT) and FAIR - The Choice Lab, Norwegian School of Economics, Bergen, Norway; emails: Alexander.Cappelen@nhh.no, Odd.Fjeldstad@cmi.no, Mmari@repoa.or.tz, Ingrid.Sjursen@snf.no and Bertil.Tungodden@nhh.no.

1 Introduction

A comprehensive economic and political economy literature documents the fact that resource-rich countries often perform poorly in terms of social and economic development compared with countries that are less abundant in natural resources (Ross 2015; Venables 2016). Natural resource wealth is associated with less democracy (Aslaksen 2007; Andersen and Aslaksen 2013; Ramsay 2011), more corruption (Leite and Weidmann 2002; Arezki and Brückner 2011; Vicente 2010; Sala-i-Martin and Subramanian 2013; Andersen et al. 2013; Brollo et al. 2013; Caselli and Michaels 2013) and a higher likelihood of violent conflicts (Collier and Hoeffler 1998; Ross 2004). These phenomena are commonly referred to as the *the resource curse*, a term first proposed by Auty (1993). Although there seems to be a consensus that a resource curse often exists, the empirical literature faces greater challenges in establishing *why* natural resource wealth is often associated with undesirable outcomes, because cross-country comparisons are plagued with endogeneity issues (Cust and Poelhekke 2015).

The motivation of this paper is to take an alternative approach and study one possible micro-founded mechanism for the resource curse, using a controlled field experimental setting in which we focus on corruption, which is often considered to be a key explanation of the resource curse (Torvik 2002; Svensson 2005; Serra 2006; Campos and Pradhan 2007; Kolstad and Wiig 2009; Olken and Pande 2012). The mechanism that we have in mind is that expectations about natural resource rents in a country cause people to expect an increase in corruption in the future, which in turn increases their willingness to engage in corrupt activities. To study this mechanism, we take advantage of the fact that Tanzania has recently discovered large reservoirs of offshore natural gas, but has not yet commenced production or revenue generation.[1] In this context, we implement a controlled field experiment, in which respondents are randomized to watch different versions of an *informational* video to create exogenous variations in expectations about future gas revenues, and they then respond to survey questions and take part in a behavioral experiment. This design allows us to causally identify how expectations about natural resource wealth shape people's expectations about corruption and their willingness to engage in corrupt behavior (Ross 2012; 2015).

The large-scale study, involving about 3,000 respondents, was conducted in the commercial capital of Tanzania, Dar es Salaam, and in the two gas regions, Mtwara and Lindi, in the south-eastern part of the country. Our main result is that we find support for the hypothesis that expectations about a future natural resource rents increases expectations about future corruption. The effect is not large, about 0.07 standard deviations, but it appears across almost all sub-groups in our sample and is robust to the inclusion of a large set of controls. Given that our experimental manipulation is subtle and cleanly identifies the effect of providing more information (where possible confounds are removed by the fact that the control group watched almost the same video), we argue that the result points to an important mechanism that may contribute to explaining the resource curse. An interesting part of our analysis, which was not pre-specified, is that we find a

[1] These large reservoirs of natural gas have been discovered off the southern coast of Tanzania from 2010 (URoT 2013a). The size of the total confirmed gas reserves is currently standing at more than 57 trillion cubic feet and it is expected to increase further in the years to come. International oil companies have already invested heavily in the exploration phase: the Tanzania Petroleum Development Corporation estimates that overall investments amount to almost USD 5 billion so far. A final decision about investments in extraction facilities and liquefied natural gas infrastructure to facilitate gas export has not yet been made.

particularly strong effect of the information on expectations about future corruption for older respondents in Dar es Salaam. We suggest that this finding is explained by these respondents being more likely to have experienced the mining boom and the associated increase in corruption in Tanzania in the 1990s (Lange 2011), and therefore, they update their beliefs about future corruption based on this personal experience, in line with the learning mechanism identified in Malmendier and Nagel (2016).

In contrast to our pre-specified hypothesis, however, we do not find a corresponding effect on the willingness to engage in corrupt activities, both in terms of the survey measure and the behavioral measure of the corruption norm. The survey measure involves responses to a question on the appropriateness of engaging in corruption, the majority of which express a strong norm against corruption. The behavioral measure is based on an incentivized experiment where the respondent self-reports the number of tails from six coin flips. In contrast to the survey measure, we observe extensive dishonest behavior among the respondents, particularly among the males. In both cases, however, the measures are not affected by our treatment manipulation. A possible explanation for this finding is that it takes some time for a change in beliefs to translate into a change in behavior, but it may also be explained by corrupt behavior being affected more by the present level of corruption in society than by expectations about future corruption. Overall, therefore, we find only limited evidence for our specific hypothesis that expectations about a natural resource rents contribute to the resource curse by increasing expectations about corruption in society in the future, which, consequently, increases people's present willingness to engage in corrupt activities.

Our study relates to a growing literature that focuses on political and institutional explanations of the resource curse. A central theme in this literature is that natural resource booms affect the incentives and behavior of both the political elite and ordinary citizens (Treisman 2000; Leite and Weidmann 2002; Andersen and Aslaksen 2013; Brollo et al. 2013; Caselli and Michaels 2013; Paler 2013; Cust and Poelhekke 2015) and that the extent to which agents can act on bad incentives depends on the quality of institutions (Mehlum et al. 2006; Bhattacharyya and Hodler 2010). On the government side, the argument is that natural resource revenues, particularly those arising from petroleum, increase the value to the ruling politicians of staying in power, because it gives them direct access to large rents. At the same time, these petroleum rents provide an opportunity for the elite to spend money on activities that aim to increase their chances of staying in power. Examples of such activities are increased patronage spending (for instance, increased employment in the public sector), vote buying and reduced non-resource taxation (Robinson et al. 2006; Ross 2008; 2012; Andersen and Aslaksen 2013). However, the problem of corruption and rent-seeking is not confined to the political elite, but is also likely to apply to ordinary citizens. When the rents available for grabbing are large, it becomes more profitable for individuals to engage in political lobbying and corruption to appropriate a share of the wealth (Mehlum et al. 2006; Busse and Gröning 2013). We contribute to this literature by establishing causally that expectations about natural resource rents, which are likely to make corruption more profitable, increase expectations about future corruption in a society with weak institutions and a history of extensive corruption (Lange 2011; Gray 2015).[2] This finding is in line with Vicente (2010), who, in a natural experiment comparing Sao Tome and Principe to Cabo Verde, finds that the announcement of oil discoveries increased perceived corruption in public services.

[2]Tanzania ranked in the lowest 28^{th} - 42^{th} percentile of all six dimensions of the World Governance Indicators in 2016 (World Bank 2018).

The paper also relates to a large experimental literature studying lying in settings without strategic interactions. In a meta-analysis of 90 studies, involving more than 44,000 respondents from 47 countries, Abeler et al. (2017) find that on average, people are willing to forgo 75% of a potential monetary pay-off to avoid lying. The majority of these studies are conducted on students, or other specific population groups, and in Europe and the US, although with some notable exceptions (Banerjee et al. 2016; Heldring 2016; Lowes et al. 2017). Thus, the behavioral experiment we conduct as part of this study represents the largest study of dishonest behavior in a development context, where the respondents are largely representative of their society. We observe that our respondents are slightly less honest than the average respondent in Abeler et al. (2017); in our sample, the respondents forgo about 67% of a potential monetary pay-off to avoid lying. However, comparing our results to previous studies from Tanzania, our sample is more honest (Abeler et al. 2017; Di Falco et al. 2016). We also find large heterogeneities in dishonest behavior between sub-groups of the sample, which highlights the importance of moving beyond student samples in studies of dishonest behavior. Finally, the paper contributes to a growing body of studies using variation in information provided to respondents in surveys to identify causal effects on beliefs and preferences (Jensen 2010; Card et al. 2012; Cruces et al. 2013; Kuziemko et al. 2015).

The rest of the paper is structured as follows. Section 2 describes the experimental design and provides an overview of the sample and the data. Section 3 explains the empirical strategy and Section 4 provides a descriptive analysis of the key outcomes of interest and reports on main analysis, regarding the effect of providing information about gas revenues on expected corruption. Finally, Section 5 discusses the results and their policy implications and concludes the paper.

2 Sample and experimental design

In this section, we provide a discussion of the sample and the experimental design of the survey.

2.1 Sample

The study was conducted in July and August, 2015, in three regions in Tanzania: Dar es Salaam, Lindi and Mtwara. Dar es Salaam was chosen because it is the commercial capital. Mtwara and Lindi were chosen because they are the regions closest to the offshore gas reservoirs, and they are also the regions where a planned liquefied natural gas processing plant will be located. In the analysis, we refer to Dar es Salaam as the "non-gas region" and Mtwara and Lindi as the "gas regions".[3]

In each region, three districts/municipalities were selected. All municipalities in Dar es Salaam were included, as it possesses three districts. Both Lindi and Mtwara have more than three districts and thus, we undertook a sampling of districts in these regions. In Lindi, we selected the only urban district in the region to ensure urban representation, and randomly selected two of the rural districts. In Mtwara, we selected the urban district situated along the coast because it is the closest to the gas reservoirs, and randomly

[3]The geographical locations of the three regions are shown in Figure A.1 in Appendix A. In Tables A.3, A.4 and A.5, we provide the main analysis separately for each of the three regions.

selected two rural districts.[4] Within each of the nine districts, three wards were randomly selected using data from the 2012 Population and Household Census (URoT 2016a;c;b). In the field, three villages/streets within each of the 27 wards were randomly selected and within each of the approximately 81 villages/streets, between 35 and 40 households were randomly interviewed. In total, we surveyed 3,004 households. From each household, we randomly selected one person above 18 years of age and alternated between interviewing a man and a woman.

Table 1 provides an overview of the respondents' background variables for the sample as a whole, as well as for each of the three regions, together with comparable statistics from the 2012 Population and Household Census (URoT 2013b; 2015; 2016a;c;b). A direct comparison with the national census data is difficult in relation to age, education and marital status because our sample covers only those aged 18 years and above, whereas in the national census, the education level is measured for all individuals above five years old, and marital status is recorded for individuals above 15 years old. We observe, however, that our data is comparable to the national census in terms of occupation and gender. In the comparison between the non-gas region and the gas regions, we note that the main differences relate to education levels and occupations: respondents in the non- gas region are much more likely than those in the gas region to have completed higher education and to be self-employed.

[Table 1 about here.]

2.2 Survey and experimental design

First, we provide an overview of the structure of the survey, before turning to a more detailed discussion of the treatment variation and the main outcome variables.

General structure. The survey was conducted in Swahili, the official language of Tanzania. It lasted approximately 15 minutes and consisted of five parts (see Appendix B.1 for the English wording of the survey). In the first part, respondents answered background questions about their age, marital status, region of origin, region they visit most frequently, education and occupation. The second part entailed the experimental part of our research design: respondents were randomly assigned to watch one of two versions of an informational video on the enumerator's tablet, where one version contained information about gas discoveries and revenue estimates for Tanzania, whereas the other did not mention gas revenues at all. In the third part, we asked our pre-specified main outcome questions on corruption, trust and taxation, and additional questions that could provide a greater understanding of the mechanisms driving the responses to these questions.[5] In the fourth part, respondents were randomized to take part in one of two incentivized experiments, intended to measure behaviorally their willingness to engage in corrupt activities and their trust level. The respondents were paid upon completion of the experimental task, and the interview was then terminated. At the end, the enumerator recorded information about the gender of the respondent, the region, district, ward and

[4]The Dar es Salaam districts at the time of the study were Ilala, Temeke and Kinondoni. The Mtwara districts selected were Masasi Vijijini, Mtwara Manisipaa and Newala, and the Lindi districts were Lindi, Lindi Manisipaa and Nachingwea.

[5]Even though corruption, taxation and trust are related topics, we consider that they raise different questions and relate to different literatures. Therefore, we focus on corruption in the present paper; taxation and trust will be analyzed in separate papers.

village where the interview was conducted. The data were collected using tablets and the Qualtrics Offline Surveys app, allowing us to randomize the treatment at the enumerator level.[6] A schematic representation of the research design is provided in Figure 1 below.

[Figure 1 about here.]

The treatment. The respondents were randomly assigned to watch an informational video with or without gas information (see Figures B.1 - B.14 in Appendix B.2 for screenshots).[7] To make the comparison between the two groups as clean as possible, the video shown to the no gas information group was also part of the version providing gas information.

The no gas information sequence contains only general information about Tanzania. It describes the size of the population (49 million), the number of regions (30) and the gross national income (81 trillion Tanzanian shillings (TZS)/USD 41 billion) of the country. Importantly, this general information features components displayed in the gas information sequence that could affect expectations about corruption, such as a map of Tanzania and its regions, the Tanzanian flag, a picture of a TZS 10,000 note and the mention of a large sum of money. Because these components are shown to respondents in both the gas and the no gas information groups, we assume that they cannot explain any observed treatment differences.

The gas information video also contains a sequence explaining that natural gas has been discovered in Tanzania and indicating the location of the offshore reservoirs, off the coast of Mtwara and Lindi. Further, it describes that the gas can be extracted and sold, and that it can yield an estimated TZS 106 trillion in total revenue.[8]

The videos were made in collaboration with DJPA Tanzania and contained simple animations as well as written text and a voice-over in Swahili. The respondents watched the video on the enumerator's tablet, with a headset to hear the voice-over. Before the main data collection, the videos were tested in focus group discussions with Dar es Salaam residents to ensure that they were clear and understandable.

Overall, the experimental design creates exogenous variation between respondents regarding whether they were informed about the gas revenues, before they provide answers on the main outcome variables. The presence of a control video ensures that there is no variation in other components. Thus, the present design provides for a clean identification of the effect of being informed about the likelihood of natural resources yielding large revenues to Tanzania in the future. Even though our treatment design is subtle, and one should not expect it to generate large effects, we believe that it may trigger two mechanisms that could affect how the respondents answer the questions that follow. First, for respondents who are aware of the potential gas revenues, the information may cause

[6] As shown in Table A.1, the sample is balanced both for all respondents and within the gas region and non-gas regions.

[7] As a robustness check, we introduced four variations of this video in which we randomly varied how we illustrated the size of the revenue estimates. Specifically, we varied whether the revenue estimates were presented in terms of total value/annual real return and in terms of all Tanzanians/per capita. As shown in Tables A.8, A.9 and A.10 in Appendix A, we do not find significant differences across these sub-treatments. We aimed to allocate an equal share of the respondents to each of the five versions of the video, but a software problem created some deviations. Overall, as shown in Table A.2, around 23% of the respondents were in the control group with no gas information and about 77% of the respondents were in the treatment group with gas information.

[8] This estimated total value is based on IMF (2014).

them to update the value of these resource rents upwards or downwards. Second, for all respondents, the information about gas revenues increases the salience of this revenue. To shed some light on the extent to which these mechanisms shape the following responses, immediately after the video we asked all respondents whether the information was new to them, as follows:

- *Novelty of information*: How much of this information was new to you? (All of it - most of it - some of it - almost none of it - none of it)

In addition, for the treated respondents, we asked whether the estimated total gas revenues were larger than they had expected.

The main survey outcome variables. We were interested in whether information about future gas revenues causes people to expect more corruption in society and weaken the social norm whereby corruption is considered to be unacceptable. To study these two dimensions, we asked the respondents two main questions on corruption after they had seen the informational video:

- *Expected future corruption*: In the years to come, I expect the extent of corrupt activities to: (Decrease a lot - decrease - stay the same - increase - increase a lot)

- *Corruption norm*: People should never engage in corrupt activities (Strongly agree - agree - neither agree nor disagree - disagree - strongly agree)

Our prior expectation was that information about future gas revenues would cause individuals to expect more corruption if they updated their estimates of the resource rents upwards. Further, we expected this information to make the presence of a large resource rents salient, which might remind respondents of the extensive corruption involved in historical cases of national resource rents extraction in Tanzania (Lange 2011). Finally, building on the prior expectation that the information video would make respondents expect more corruption, we also expected the information to weaken the corruption norm because, as well established in the literature, the strength of a corruption norm depends importantly on the level of corruption in society (Andvig and Moene 1990; Treisman 2000; Paldam 2002; Fisman et al. 2007; Cameron et al. 2009; Barr and Serra 2010).

The behavioral outcome variable. To supplement the survey response on the corruption norm, we conducted an incentivized experiment to measure people's dishonesty. To capture this aspect, we conducted a standard flip-of-a-coin experiment (Fischbacher and Föllmi-Heusi 2013), where the respondent was asked to flip a coin six times without the enumerator observing the outcomes. Then, the respondent was asked to report how many heads he/she obtained, with a payment of $1,000$ TZS per tail reported. The payment scheme was announced before the respondent flipped the coin. With this experimental design, we placed the participant in a situation where he or she could "abuse entrusted power for private gain", which is a standard definition of corrupt behavior. We thus take the number of tails reported as a group level measure of dishonesty in our study.[9]

The idea behind this experimental task was that the enumerator entrusted power to the respondent to report the correct outcomes, but the respondent could, without any

[9]Corruption is a multifaceted concept (Svensson 2005; Campos and Pradhan 2007; Mishra 2018), but dishonesty is an essential feature of any notion of corrupt behavior (Gächter and Schulz 2016; Gneezy et al. Forthcoming).

fear of detection, abuse this power to benefit economically by misreporting the number of heads and tails. Our experimental design allows us to identify whether information about future gas revenues makes it more likely that the respondent abuses their power and engages in corrupt activity for a private gain.

3 Empirical strategy

We study the causal effect of providing information about gas on the respondents' responses by estimating the following linear ordinary least squares regression, as specified in the pre-analysis plan:[10]

$$y_i = \alpha + \beta_{GI} GI_i + \delta X_i + \varepsilon_i, \tag{1}$$

where y_i is the standardized version of the relevant outcome measure for individual i (expected corruption, corruption norm and novelty of information), GI_i is a treatment indicator variable that takes a value of one if respondent i was exposed to the gas information version of the video and X_i is a vector of individual characteristics: age, gender, occupation, education, marital status and type of region (gas or non-gas region). We estimate Equation (1) with robust standard errors and control for round fixed effects.[11] We also report the corresponding regressions without control variables. As respondents were randomly assigned to watch one of the two versions of the video, β_{GI} provides an estimate of the causal effect of providing information about gas on expectations about future corruption and on the corruption norm.

To investigate whether the gas information affected sub-groups of the sample differently, we also estimate regressions with interaction terms. We focus on age, education, gender and employment status, and estimate the following regression with indicator variables for the respective dimensions:[12]

$$y_i = \alpha + \beta_{GI} GI_i + \gamma Var_i + \theta GI_i \times Var_i + \delta X_i + \varepsilon_i, \tag{2}$$

where Var_i is an indicator variable taking a value of one if respondent i is older than the median age, has completed lower secondary school or a higher level of education, is male, is self-employed or is a farmer, and $GI_i \times Var_i$ is an interaction term between GI_i and Var_i. Then, the estimated effect on respondents of providing information about gas is given by β_{GI} (younger, less educated, females, employed in the formal sector) and $\beta_{GI} + \theta$ (older, more educated, males, self-employed, farmers) and the estimated difference in causal effect between the two respective sub-groups is given by θ.

[10] The pre-analysis plan was registered with the American Economic Association's registry for randomized controlled trials: AEARCTR-0000768.

[11] The round fixed effects are included because, initially, owing to a technical problem, we allocated too many respondents to the treatment group. To address this problem and ensure that we reached the planned targets for the different groups, we divided the data collection into five rounds. In the first round, 90% of respondents were allocated to the gas information video, whereas in rounds 2–5, the corresponding share was between 69% and 72%.

[12] Although we did not pre-specify the age dimension in our pre-analysis plan, we included it because we found some interesting patterns that fit well with the findings of related literature.

4 Analysis

First, we provide some descriptive statistics of the responses of the respondents before turning to the main analysis of the average treatment effects and heterogeneous effects on sub-groups in our sample.

4.1 Descriptive statistics

Here, we provide an overview of how the respondents answered the main questions of interest.

First, as shown in Figure 2, we observe that the majority of respondents found that the information video provided some new information (85% of respondents answered that some, almost all or all information was new to them). Further, we observe that this share is slightly larger in the gas regions than in the non-gas region (87% versus 81%, respectively, $p = 0.000$), which is in line with expectations, given that the level of education is higher in the non-gas region than in the gas regions.

[Figure 2 about here.]

Further, to shed light on how the gas information affected respondents in the treatment group, we asked them whether the estimated total gas revenues presented in the video were larger than expected. For the overall sample, we find that 65% of the sample did so, which provides evidence that, on average, the information caused the treated respondents to update the natural resource rents upwards. As shown in Figure 3, the pattern applies across all pre-specified sub-samples (with the majority in each sub-group answering "Yes" to this question). In particular, we note that the less educated are much more likely to find the estimated total revenues larger than expected.

[Figure 3 about here.]

Figure 4 shows the distribution of expected future corruption in the sample. We note that there is large variation in the respondents' beliefs. The majority (60%) expect corruption to increase or increase a lot in the years to come, but a significant share of the respondents (30%) expect a decrease in corruption. The patterns are not significantly different for the gas regions and non-gas region.

[Figure 4 about here.]

In Figure 5, we present an overview of the responses on both the survey measure and the behavioral measure of the corruption norm. In terms of the survey measure, we observe a strong norm against engaging in corruption, with about 60% of respondents strongly agreeing with this norm in both the gas and non-gas regions. In contrast, we observe that a significant share of the respondents are willing to engage in a corrupt activity in the behavioral experiment.[13] In the overall sample, on average, the respondents report four tails, which is significantly larger than the benchmark of three tails

[13]We also asked them the non-incentivized question of how likely they believed it was that others would engage in corrupt behavior. The responses to this question are in line with what we observe in the coin flipping task.

that would be the expected outcome from honest reporting.[14] We observe that dishonest reporting is more prevalent in the non-gas region than in the gas regions (4.6 tails versus 3.7 tails, $p = 0.000$) and, in particular, we note that among males in the non-gas region, the average number of reported tails is very high, at 4.8.[15]

[Figure 5 about here.]

4.2 Treatment analysis

In this part of the paper, we analyze how providing information about gas revenues causally affected the respondents' expectations about future corruption and the corruption norm.

Table 2 reports ordinary least squares regressions for which the dependent variable is expectations about future corruption. In line with our pre-specified hypothesis, as shown in column (1), we find a significant positive causal effect of providing information about gas revenue on expected future corruption, equal to about 0.07 standard deviations (two-sided test, $p = 0.080$). In addition, in column (2), we observe that this effect is robust to the inclusion of control variables. In columns (3) and (4), we report the corresponding analysis separately for the non-gas regions and the gas region. In both regressions, we observe the same positive effect (even though the point estimates are now less precise, owing to a smaller number of observations). In terms of control variables, we observe systematically that males expect more corruption in the future than do females.

[Table 2 about here].

When evaluating the size of the estimated causal effect on beliefs about future corruption, it is important to keep in mind that the experimental manipulation is subtle in order to cleanly identify the relationship between beliefs about the resource rents and beliefs about corruption. In the informational video shown to the treatment group, we only added information about the estimated total gas revenues. Our first main finding is that this information caused the respondents to expect more corruption in the future.

Result 1: We find a significant causal effect of providing information about the estimated size of the natural resource rents on beliefs about corruption in the future.

In Table 3, we report regressions for the estimated effect of providing information about gas revenues on the corruption norm, as measured by the survey question and the behavioral experiment. Surprisingly, given the positive effect on expectations about future corruption, we do not find any effect on the corruption norm. As shown in columns (1) and (4), we estimate a relatively precise null effect for the whole sample on both the

[14] Compared to the average in the meta analysis conducted by Abeler et al. (2017), our sample is more dishonest (our sample was willing to forgo 67% of the potential monetary pay-off to avoid lying, whereas the corresponding percentage in the meta-analysis was 75%). However, compared with other studies from Tanzania (Abeler et al. 2017; Di Falco et al. 2016), we find relatively less dishonest behavior.

[15] Consistent with Abeler et al. (2017), we also find that, overall, men are more dishonest than women (4.3 tails versus 3.6 tails, $p = 0.000$) and younger respondents are more dishonest than older respondents (4.4 tails versus 3.6 tails, $p = 0.000$).

survey measure and the behavioral measure. This also occurs when we run the regressions separately for the non-gas region and the gas regions (columns (2)–(3) and (5)–(6), respectively).[16]

[Table 3 about here.]

Result 2: We do not find a causal effect of providing information about the estimated size of the natural resource rents on people's willingness to endorse or engage in corrupt activities.

The null effect on the corruption norm may reflect that the acceptance of and willingness to engage in corrupt activities is related more to the present level of corruption in society than to expectations about future corruption.

To shed further light on the underlying mechanism driving the causal effect of providing information about total revenues on expected corruption, in Table 4 we investigate whether providing information about gas revenues affected the perceived novelty of the informational video. Surprisingly, as shown in columns (1) and (2), although we do not find a treatment effect on novelty for the whole sample, this null effect hides two opposing patterns in the non-gas region and in the gas regions, as shown in columns (3) and (4). In the non-gas region, we find the expected strong positive effect of providing information on perceived novelty, whereas we find a significant negative effect in the gas regions. A possible interpretation of the negative effect in the gas regions is that the local population in this area is quite well informed about the possible revenues from the gas fields, given their location closer to the fields and the prominence of this issue in the local debates. As a result, the informational video may have appeared less novel when this sequence was added to the video. In contrast, the potential for revenues from gas-field production has figured less prominently in the non-gas region and, thus, the sequence with information about expected total revenues may appear more novel to this population.

[Table 4 about here.]

We find very different patterns in terms of how the respondents view the novelty of information in the non-gas regions and the gas region, but the same effect on expectations about corruption. This suggests that providing information on total revenues not only operates through updating the respondents' beliefs about the natural resource rents, but also through making gas production a salient issue before they answer the questions on corruption in the survey.

4.3 Heterogeneity analysis

Here, we investigate whether different sub-groups of the population are affected differently by the gas information, with a focus on the effect on expectations of corruption.[17]

[16] We also observe the same finding if we consider the alternative survey measure on the corruption norm, see Table A.11.

[17] In Table A.6 and Table A.7, we show the corresponding heterogeneity analysis for the effect on the corruption norm. Table A.12 shows the heterogeneity analysis for an alternative measure of the corruption norm.

[Table 5 about here]

In Table 5, we report the heterogeneity analysis for the non-gas region and the gas regions separately. We observe that for almost all sub-groups, the estimated causal effect is positive, which may be seen as evidence of information about total revenues systematically causing the respondents to expect more corruption in the future. We should also note that the point estimates are not precisely measured and, for many of the groups, they are not statistically significant. However, we find a strong and highly significant effect among the older respondents in the non-gas region: the information about the total revenues causes an increase in their expectations of future corruption by about 0.3 standard deviations ($p = 0.002$). Older respondents are more likely to have experienced the mining boom and the associated increase in corruption in the 1990s in Tanzania (Lange 2011). This personal experience may have influenced how older respondents updated their beliefs about future corruption, in line with the learning mechanism observed in the context of expectations about future inflation in Malmendier and Nagel (2016). We consider the finding for the older respondents to be explorative because we did not prespecify this hypothesis, and we note that it is less strong among the older respondents in the gas regions. Nevertheless, we find it important and worthy of further investigation because it may shed light on the persistence of corruption in many developing countries.

Result 3: We find that providing information about the estimated size of the natural resource rents increases people's expectations about future corruption in most sub-groups, particularly among the older respondents in the non-gas region.

In Table 6, we report results for the heterogeneity analysis of how gas information affected the perceived novelty of information. Interestingly, we observe systematic, but opposite, patterns in the non-gas region and the gas regions. In almost all sub-groups in the non-gas region, the information has a strong positive causal effect on perceived novelty, whereas in almost all sub-groups in the gas regions, it has a strong negative effect (although not always statistically significant). Thus, these patterns strongly suggest that the information manipulation worked differently in the non-gas regions and the gas regions; in the gas regions, the effect on beliefs about future corruption primarily seems to have been driven by the natural resource issue being made salient for the respondents.

In light of Result 3, we note that the information has a strong positive effect on perceived novelty equal to about 0.3 standard deviations for the older respondents in the non-gas region, which is in line with the effect on future corruption for this group. Nevertheless, the effect on perceived novelty is not significantly stronger than for most of the other sub-groups in the non-gas region, which suggests that the particularly strong effect on beliefs about future corruption for this group operated through the mechanism of their beliefs being updated differently to those of other groups; in particular, this arises because the older respondents are more likely to have experienced the mining boom and the corresponding increase in corruption.

[Table 6 about here.]

5 Conclusions

It is important to gain a better understanding of the underlying mechanisms of the resource curse. In this paper, we report from a large-scale field experiment in Tanzania that offers causal evidence on how information about a natural resource rents shapes people's beliefs about future corruption. We introduce a clean experimental design that controls for a host of potential confounds by minimally manipulating an informational video also shown to the control group. We find that our information manipulation significantly affects the respondents' beliefs about future corruption, particularly among the older respondents. We believe that this result points to a mechanism that may play an important role in generating the resource curse: the presence of natural resource rents makes people expect more corruption and the increased expectations about corruption may become self-fulfilling by making people more willing to engage in corrupt activity (Andvig and Moene 1990). When facing potentially large future resource rents it thus becomes even more important for countries to strengthen their institutions to avoid such a negative spiral of corruption. However, in our study, we only find limited evidence for this mechanism because our treatment manipulation does not change people's willingness to engage in corrupt behavior.

We believe that several interesting research avenues arise from the present study. First, more research is needed to understand the mechanisms driving corrupt behavior, and particularly how it relates to people's beliefs about corruption in society. We do not find evidence that beliefs about future corruption shape people's normative views on corruption or their behavior, which is suggestive of corrupt behavior instead being driven by people's beliefs about the existing level of corruption in a society. Second, the strong effect of the information manipulation on older respondents is intriguing. We interpret this finding in the light of the learning framework of Malmendier and Nagel (2016), who show that belief updating is shaped by personal experiences. In our study, this may suggest that the effect on the older respondents is driven by their knowledge about the extensive corruption in the mining sector in Tanzania. This hypothesis was not pre-specified and, thus, more research is needed to understand how personal experiences of this kind may shape people's beliefs about corruption. More generally, we believe that the learning framework represents a fruitful approach for research because it introduces an interesting view on how historical experiences may shape the development process. Finally, we believe that our experimental design offers a new approach to the study of the resource curse that can be used in a variety of settings. In particular, we believe that it would be of interest to study how politicians respond to information about natural resource rents, in terms of both their beliefs about future corruption and their willingness to engage in corrupt behavior.

References

Abeler, Johannes, Daniele Nosenzo, and Collin Raymond (2017). "Preferences for truth-telling," mimeo.

Andersen, Jørgen Juel and Silje Aslaksen (2013). "Oil and political survival," *Journal of Development Economics*, 100(1): 89–106.

Andersen, Jørgen Juel, Niels Johannesen, David Dreyer Lassen, and Elena Paltseva (2013). "Petro rents, political institutions, and hidden wealth: Evidence from bank deposits in tax havens," Econ. Policy Res. Unit Work. Pap. No. 2013-03.

Andvig, Jens Chr. and Karl Ove Moene (1990). "How corruption may corrupt," *Journal of Economic Behavior and Organization*, 13(1): 63–76.

Arezki, Rabah and Markus Brückner (2011). "Oil rents, corruption, and state stability: Evidence from panel data regressions," *European Economic Review*, 55(7): 955 – 963.

Aslaksen, Silje (2007). "Corruption and oil: Evidence from panel data," Unpublished.

Auty, Richard M. (1993). *Sustaining Development in the Mineral Economies: The Resource Curse Thesis*, Routledge.

Banerjee, Ritwik, Datta Nabanita Gupta, and Marie Claire Villeval (2016). "The spillover effects of affirmative action on competitiveness and unethical behavior," IZA Discussion Paper No. 10394.

Barr, Abigail and Danila Serra (2010). "Corruption and culture: An experimental analysis," *Journal of Public Economics*, 97(11): 862–869.

Bhattacharyya, Sambit and Roland Hodler (2010). "Natural resources, democracy and corruption," *European Economic Review*, 54(4): 608 – 621.

Brollo, Fernanda, Tommaso Nannicini, Roberto Perotti, and Guido Tabellini (2013). "The political resource curse," *American Economic Review*, 103(5): 1759–96.

Busse, Matthias and Steffen Gröning (2013). "The resource curse: Governance and resources," *Public Choice*, 154(1-2): 1–20.

Cameron, Lisa, Chaudhuri Chaudhuri, Nisvan Erkal, and Lata Gangadharan (2009). "Propensities to engage in and punish corrupt behavior: Experimental evidence from Australia, India, Indonesia, and Singapore," *Journal of Public Economics*, 93(7): 843–852.

Campos, J. Edgardo and Sanjay Pradhan (2007). *The many faces of corruption: Tracking vulnerabilities at the sector level*, World Bank.

Card, David, Alexandre Mas, Moretti Enrico, and Saez Emmanuel (2012). "Inequality at work: The effect of peer salaries on job satisfaction," *American Economic Review*, 102(6): 2981–3003.

Caselli, Francesco and Guy Michaels (2013). "Do oil windfalls improve living standards? Evidence from Brazil," *American Economic Journal: Applied Economics*, 5(1): 208–238.

Collier, Paul and Anke Hoeffler (1998). "On economic causes of civil war," *Oxford Economic Papers*, 50(4): 563–573.

Cruces, Guillermo, Ricardo Perez-Truglia, and Martin Tetaz (2013). "Biased perceptions of income distribution and preferences for redistribution: Evidence from a survey experiment," *Journal of Public Economics*, 98: 100–112.

Cust, James and Steven Poelhekke (2015). "The local economic impacts of natural resource extraction," OxCarre Research Paper 156.

Di Falco, Salvatore, Brice Magdalou, David Masclet, Marie Claire Villeval, and Marc Willinger (2016). "Can transparency of information reduce embezzlement? experimental evidence from tanzania," IZA Discussion Papers, No. 9925.

Fischbacher, Urs and Franziska Föllmi-Heusi (2013). "Lies in disguise – An experimental study on cheating," *Journal of the European Economic Association*, 11(3): 525–547.

Fisman, Raymond J., Shachar Kariv, and Daniel Markovits (2007). "Individual preferences for giving," *American Economic Review*, 97(5): 1858–1876.

Gächter, Simon and Jonathan F Schulz (2016). "Intrinsic honesty and the prevalence of rule violations across societies," *Nature*, 531: 496–499.

Gneezy, U., S. Saccardo, and R. Van Veldhuizen (Forthcoming). "Bribery: Behavioral drivers of distorted decisions," *Journal of the European Economic Association*.

Gray, Hazel S. (2015). "The political economy of grand corruption in Tanzania," *African Affairs*, 114(456): 382–403.

Heldring, Leander (2016). "Violence and the state: Evidence from Rwandas decade of atrocities," mimeo, Harvard University.

IMF (2014). "United Republic of Tanzania, selected isssues," IMF Country Report 14/22, International Monetary Fund.

Jensen, Robert (2010). "The (perceived) returns to education and the demand for schooling," *Quarterly Journal of Economics*, 125(2): 515–548.

Kolstad, Ivar and Arne Wiig (2009). "Is transparency the key to reducing corruption in resource-rich countries?" *World Development*, 37(3): 521 – 532.

Kuziemko, Ilyana, Michael I. Norton, Emmanuel Saez, and Stefanie Stantcheva (2015). "How elastic are preferences for redistribution? Evidence from randomized survey experiments," *American Economic Review*, 105(4): 1478–1508.

Lange, Siri (2011). "Gold and governance: Legal injustices and lost opportunities in Tanzania," *African Affairs*, 110(439): 233–252.

Leite, Carlos and Jens Weidmann (2002). "Does mother nature corrupt? Natural resources, corruption, and economic growth," in George Abed and Sanjeev Gupta (eds.), "Governance, corruption and economic performance," International Monetary Fund.

Lowes, Sara, Nathan Nunn, James A. Robinson, and Jonathan L. Weigel (2017). "The evolution of culture and institutions; Evidence from the cuba kingdom," *Econometrica*, 85(4): 1065–1091.

Malmendier, Ulrike and Stefan Nagel (2016). "Learning from inflation experiences," *Quarterly Journal of Economics*, 131(1): 53–87.

Mehlum, Halvor, Karl Ove Moene, and Ragnar Torvik (2006). "Institutions and the resource curse," *Economic Journal*, 116(508): 1–20.

Mishra, Ajit (2018). *Markets, Governance, and Institutions in the Process of Economic Development*, chapter The Many Faces of Corruption: Which one(s) to target?, Oxford University Press.

Olken, Benjamin A and Rohini Pande (2012). "Corruption in developing countries," *Annu. Rev. Econ.*, 4(1): 479–509.

Paldam, Martin (2002). "The cross-country pattern of corruption: Economics, culture and the seesaw dynamics," *European Journal of Political Economy*.

Paler, Laura (2013). "Keeping the public purse: An experiment in windfalls, taxes, and the incentives to restrain government," *American Political Science Review*, 107(4): 706–725.

Ramsay, Kristopher W. (2011). "Revisiting the resource curse: natural disasters, the price of oil, and democracy," *International Organization*, 65(03): 507–529.

Robinson, James A., Ragnar Torvik, and Thierry Verdier (2006). "Political foundations of the resource curse," *Journal of Development Economics*, 7(2): 447–468.

Ross, Michael (2008). "But seriously: does oil really hinder democracy?" mimeo, University of California, Los Angeles.

Ross, Michael L. (2004). "What do we know about natural resources and civil war?" *Journal of Peace Research*, 41(3): 337–56.

Ross, Michael L. (2012). *The oil curse: how petroleum wealth shapes the development of nations*, Princeton University Press.

Ross, Michael L. (2015). "What have we learned about the resource curse?" *Annual Review of Political Science*, 18: 239–59.

Sala-i-Martin, Xavier and Arvind Subramanian (2013). "Addressing the natural resource curse: An illustration from Nigeria," *Journal of African Economies*, 22(4): 570–615.

Serra, Danila (2006). "Empirical determinants of corruption: A sensitivity analysis," *Public Choice*, 126: 225–256.

Svensson, Jakob (2005). "Eight questions about corruption," *Journal of Economic Perspectives*, 19(3): 19–42.

Torvik, Ragnar (2002). "Natural resources, rent seeking and welfare," *Journal of Development Economics*, 67(2): 455 – 470.

Treisman, Daniel (2000). "The causes of corruption: a cross-national study," *Journal of Public Economics*, 76(3): 399 – 457.

URoT (2013a). "The national natural gas policy of Tanzania," The United Republic of Tanzania.

URoT (2013b). "Population distribution by age and sex," The United Republic of Tanzania.

URoT (2015). "Literacy and education monograph," The United Republic of Tanzania.

URoT (2016a). "Basic demographic and socio-economic profile, Dar es Salaam region," The United Republic of Tanzania.

URoT (2016b). "Basic demographic and socio-economic profile, Mtwara region," The United Republic of Tanzania.

URoT (2016c). "Basic dempgraphic and socio-economic profile, Lindi region," The United Republic of Tanzania.

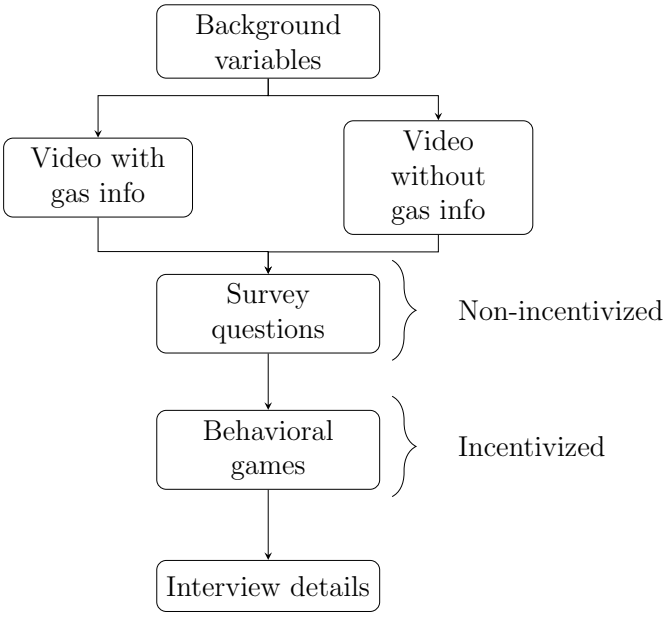
Venables, Anthony J. (2016). "Using natural resources for development: Why has it proven so difficult?" *Journal of Economic Perspectives*, 30(1): 161–84.

Vicente, Pedro C. (2010). "Does oil corrupt? evidence from a natural experiment in west africa," *Journal of Development Economics*, 92(1): 28 – 38.

World Bank (2018). "World Governance Indicators," Online dataset.

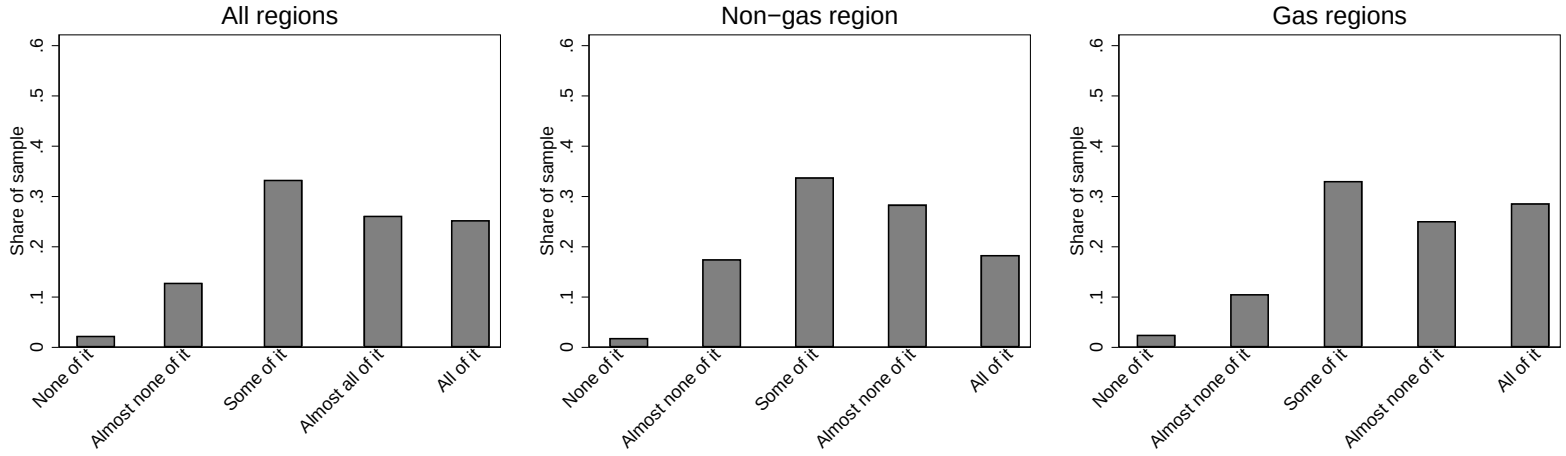
Figures and tables

Figure 1: Research design



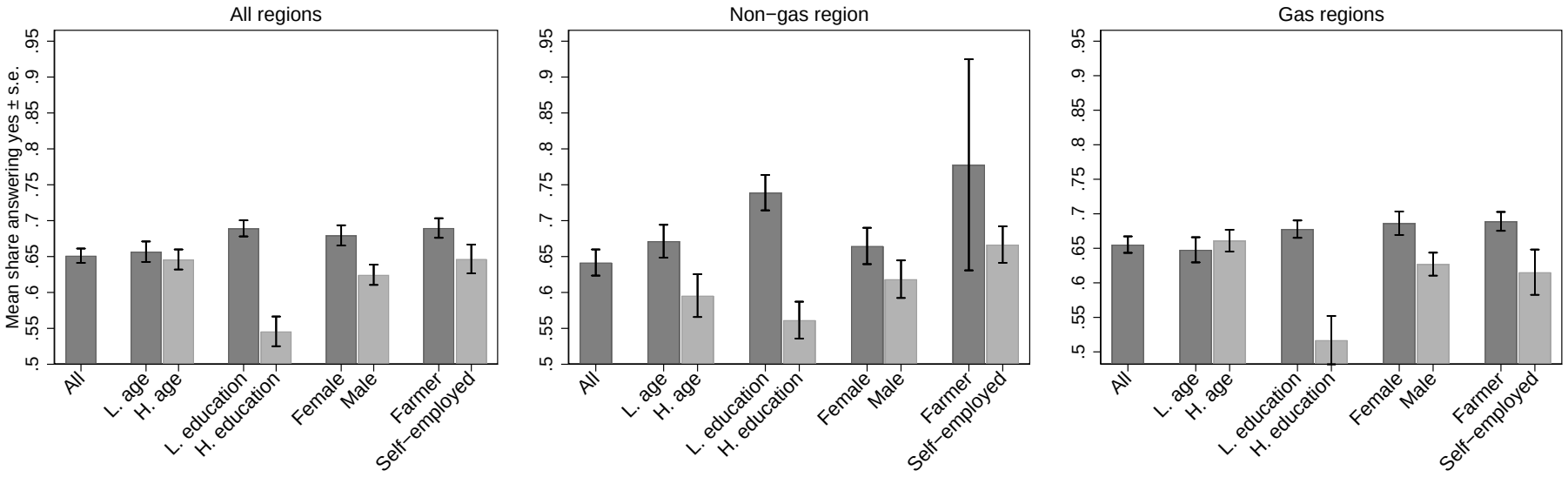
Note: The figure illustrates the sequence of the survey.

Figure 2: Perceived novelty of information, by total, non-gas and gas regions



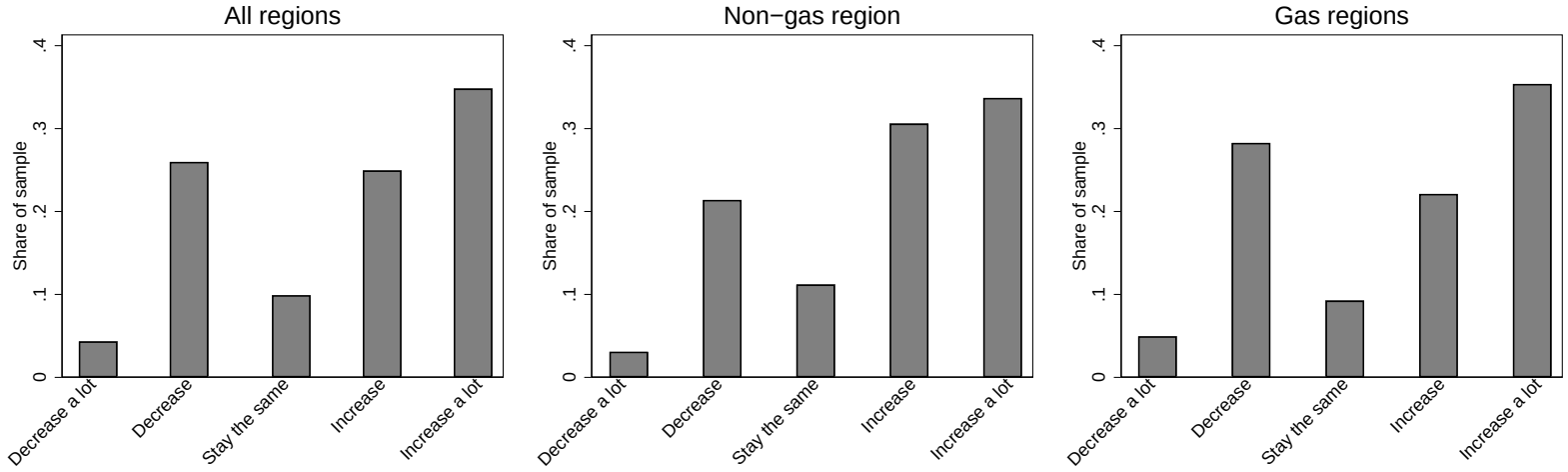
Note: The figure illustrates the distribution of responses to the question "How much of this information was new to you?" for all regions (left panel), the non-gas region (middle panel) and the gas regions (right panel).

Figure 3: Did you find the estimated total gas revenues to be larger than expected?



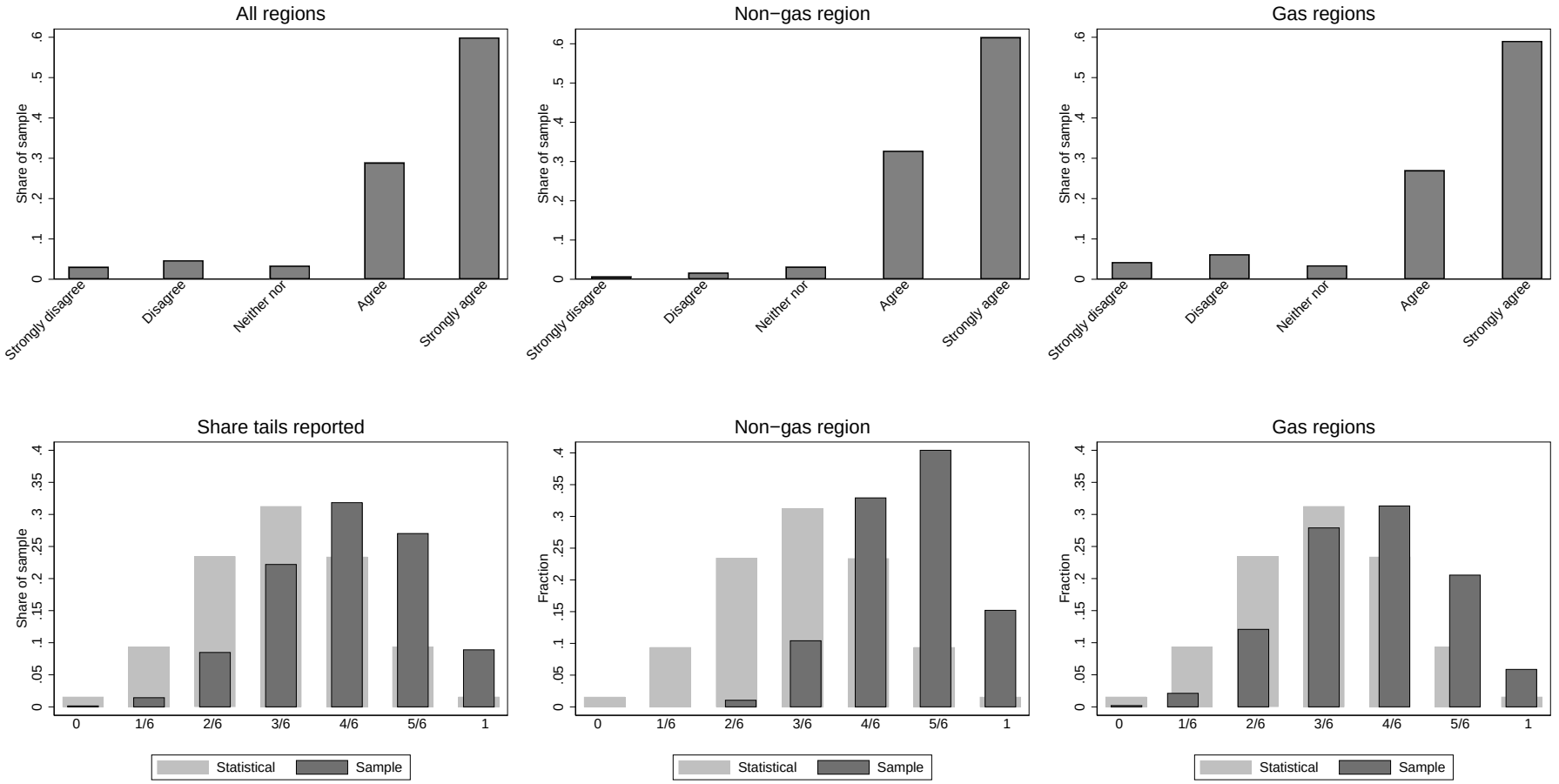
Note: The figure illustrates the share of respondents answering "yes" to the question "Did you find the estimated total gas revenue larger than you had expected?" for all sub-groups, for younger vs. older respondents, for less vs. more educated respondents, for females vs. male respondents and for farmers vs. self-employed respondents. The left panel shows results for the total sample, the middle panel shows results for the non-gas region and the right panel shows results for the gas regions.

Figure 4: Distribution of expectations about future corruption by total, non-gas and gas regions



Note: The figure illustrates the distribution of responses to the question "In the years to come, I expect the extent of corrupt activities to: decrease a lot - decrease - stay the same - increase - increase a lot" for the total sample (left panel), the non-gas region (middle panel) and the gas regions (right panel).

Figure 5: Distribution of corruption norm, by total, non-gas and gas regions



Note: **Upper panel:** The figure illustrates the distribution of responses to the question "People should never engage in corrupt activities" for the total sample (left panel), for the non-gas region (middle panel) and the gas regions (right panel). **Lower panel:** The figure illustrates the distribution of the share of tails reported in the coin flipping task (dark grey) and the statistically expected distribution for the share of tails flipped in the task (light grey) for the total sample (left panel), for the non-gas region (middle panel) and the gas regions (right panel).

Table 1: Background characteristics by region for sample and census data

	Total		Non-gas region		Gas regions	
	Sample	Census	Sample	Census	Sample	Census
Median age	34		32		36	
Higher education	0.26	0.19	0.53	0.34	0.13	0.10
	(0.01)		(0.02)		(0.01)	
Male	0.52	0.49	0.49	0.48	0.53	53
	(0.01)		(0.02)		(0.01)	
Self-employed	0.27	0.19	0.49	0.48	0.16	0.09
	(0.01)		(0.02)		(0.01)	
Farmer	0.46	0.62	0.01	0.04	0.69	0.88
	(0.01)		(0.00)		(0.01)	
Married	0.64	0.51	0.53	0.44	0.69	0.53
	(0.01)		(0.02)		(0.01)	
Observations	3004		1001		2003	

Mean coefficients; standard error of the mean in parentheses

Sample definitions: "Median age" is the median age in the sample, "Higher education" is an indicator taking a value of one if the respondent has completed lower secondary school or a higher level of education, "Male" is an indicator variable taking a value of one if the respondent is a male,, "Self-employed" is an indicator variable taking a value of one for respondents who are employed are self-employed in the non-agricultural sector, "Farmer" is an indicator variable taking a value of one for respondents whose primary occupation is farmer, "Married" is an indicator variable taking a value of one if the respondent is married and zero otherwise. *Census definitions:* "Male" is the share of males in all age groups, "Lower sec or more" is the share of the population above five years of age who have completed lower secondary school or more, "Self-employed" is the share of individuals who are 10 years or older and self-employed in the non-agricultural sector, "Farmer" is the share of individuals 10 years or older occupied in agriculture, "Married" is the share of individuals 15 years or older who are married. The median age in Tanzania is 17.7 years.

Table 2: Effect on expected future corruption of providing information

	(1) All	(2) All	(3) Non-gas region	(4) Gas regions
Gas info	0.077*	0.070*	0.086	0.057
	(0.044)	(0.041)	(0.059)	(0.056)
Above median age		–0.024	0.020	–0.039
		(0.037)	(0.065)	(0.045)
Higher education		–0.021	–0.013	–0.016
		(0.046)	(0.061)	(0.071)
Male		0.113***	0.111**	0.110***
		(0.033)	(0.056)	(0.043)
Self-employed		0.128*	0.094	0.176
		(0.067)	(0.078)	(0.133)
Farmer		0.013	–0.090	0.036
		(0.077)	(0.247)	(0.129)
Married		–0.079**	–0.076	–0.082*
		(0.038)	(0.068)	(0.048)
Gas regions		–0.016		
		(0.075)		
Enumerator FE	No	Yes	Yes	Yes
Observations	2998	2984	995	1989
R^2	0.010	0.180	0.201	0.175

Robust standard errors are shown in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: "In the years to come, I expect the extent of corrupt activities to: decrease a lot - decrease - stay the same - increase - increase a lot", standardization of five-point scale. *Treatment variable & background variables:* "Gas info" is an indicator variable taking a value of one for respondents who were exposed to gas information, "Male" is an indicator variable taking a value of one if the respondent is a male, "Above median age" is an indicator variable taking a value of one for respondents who are older than 34 years, "Higher education" is an indicator taking a value of one if the respondent has completed lower secondary school or a higher level of education, "Self-employed" is a dummy taking a value of one for respondents who are self-employed in the non-agricultural sector (the reference category is a formal sector employee), "Farmer" is an indicator variable taking a value of one for respondents who are farmers (the reference category is formal sector employee), "Married" is an indicator taking a value of one if the respondent is married and zero otherwise and "Gas region" is an indicator variable taking a value of one for respondents residing in Mtwara and Lindi (where the reference category is Dar es Salaam). *Enumerator fixed effects (FE):* indicator variables for each of the 10 enumerators. See Table A.3 in Appendix A for separate regressions for the Mtwara and Lindi regions. Six respondents refused to answer the question about expected future corruption and, therefore, the number of observations in column (1) in the table is 2998.

Table 3: Effect on corruption norm of providing information

	Survey			Behavior		
	(1) All	(2) Non-gas region	(3) Gas regions	(4) All	(5) Non-gas region	(6) Gas regions
Gas info	–0.019	0.016	–0.043	–0.033	–0.049	–0.016
	(0.037)	(0.045)	(0.053)	(0.048)	(0.070)	(0.063)
Above median age	0.051	0.109**	0.032	–0.588***	–0.431***	–0.670***
	(0.034)	(0.052)	(0.043)	(0.045)	(0.076)	(0.055)
Lower sec or more	0.061	0.053	0.057	0.011	0.028	–0.002
	(0.041)	(0.048)	(0.070)	(0.056)	(0.069)	(0.089)
Male	–0.022	–0.024	–0.024	0.480***	0.476***	0.497***
	(0.031)	(0.042)	(0.042)	(0.043)	(0.072)	(0.055)
Self-employed	–0.085	–0.069	–0.038	0.081	0.079	0.126
	(0.053)	(0.054)	(0.118)	(0.082)	(0.095)	(0.147)
Farmer	–0.189***	–0.492	–0.137	0.005	0.150	0.033
	(0.065)	(0.301)	(0.115)	(0.095)	(0.231)	(0.142)
Married	0.091**	0.005	0.112**	0.020	0.063	–0.040
	(0.036)	(0.049)	(0.049)	(0.047)	(0.079)	(0.060)
Gas regions	0.122**			–0.140*		
	(0.061)			(0.082)		
Enumerator FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2984	994	1990	1469	479	990
R^2	0.287	0.231	0.321	0.371	0.296	0.305

Robust standard errors are in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variables: Survey = "People should never engage in corrupt activities: strongly disagree - agree - neither nor - agree - strongly agree", standardization of five-point scale and behavior = share of tails reported in the coin flipping game (0, 1/6, 2/6, 3/6, 4/6, 5/6, or 1), standardized values. *Treatment variable & background variables:* "Gas info" is an indicator variable taking a value of one for respondents who were exposed to gas information, "Male" is an indicator variable taking a value of one if the respondent is a male, "Above median age" is an indicator variable taking a value of one for respondents who are older than 34, "Higher education" is an indicator taking a value of one if the respondent has completed lower secondary school or a higher level of education, "Self-employed" is a dummy taking a value of one for respondents who are self-employed in the non-agricultural sector (the reference category is a formal sector employee), "Farmer" is an indicator variable taking a value of one for respondents who are farmers (the reference category is a formal sector employee), "Married" is an indicator taking a value of one if the respondent is married and zero otherwise and "Gas region" is an indicator variable taking a value of one for respondents residing in Mtwara and Lindi (the reference category is Dar es Salaam). *Enumerator fixed effects (FE):* indicator variables for each of the 10 enumerators. See Table A.4 in Appendix A for separate regressions for the Mtwara and Lindi regions.

Table 4: Effect of providing gas information on perceived novelty of information

	(1) All	(2) All	(3) Non-gas region	(4) Gas regions
Gas info	0.049	0.028	0.224***	–0.113*
	(0.046)	(0.042)	(0.056)	(0.058)
Above median age		–0.069*	–0.165***	–0.030
		(0.037)	(0.061)	(0.045)
Higher education		–0.437***	–0.465***	–0.430***
		(0.043)	(0.055)	(0.069)
Male		–0.094***	–0.040	–0.119***
		(0.034)	(0.051)	(0.042)
Self-employed		0.226***	0.231***	0.236**
		(0.063)	(0.072)	(0.119)
Farmer		0.276***	0.235	0.318***
		(0.075)	(0.217)	(0.120)
Married		–0.027	–0.028	–0.031
		(0.038)	(0.057)	(0.049)
Gas regions		0.096		
		(0.068)		
Enumerator FE	No	Yes	Yes	Yes
Observations	3000	2986	995	1991
R^2	0.014	0.189	0.359	0.142

Robust standard errors are in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: "How much of this information was new to you? None of it - almost none of it - some of it - almost all of it - all of it", standardized five-point scale. *Treatment variable & background variables:* "Gas info" is an indicator variable taking a value of one for respondents who were exposed to gas information, "Male" is an indicator variable taking a value of one if the respondent is a male, "Above median age" is an indicator variable taking a value of one for respondents who are older than 34, "Higher education" is an indicator taking a value of one if the respondent has completed lower secondary school or a higher level of education, "Self-employed" is a dummy taking a value of one for respondents who are self-employed in the non-agricultural sector (the reference category is a formal sector employee), "Farmer" is an indicator variable taking a value of one for respondents who are farmers (the reference category is a formal sector employee), "Married" is an indicator taking a value of one if the respondent is married and zero otherwise and "Gas region" is an indicator variable taking a value of one for respondents residing in Mtwara and Lindi (the reference category is Dar es Salaam). *Enumerator fixed effects (FE):* indicator variables for each of the 10 enumerators. For separate regressions for the Mtwara and Lindi regions, see Table A.5 in Appendix A.

Table 5: Heterogeneity in effect of gas information on expected future corruption, for non-gas and gas regions

	Non-gas region					Gas regions				
	(1) H. age	(2) H. educ	(3) Male	(4) Self-empl	(5) Farmer	(6) H. age	(7) H. educ	(8) Male	(9) Self-empl	(10) Farmer
Gas info	–0.031	0.159*	0.119	0.109	0.087	0.030	0.041	–0.029	0.057	0.182**
	(0.074)	(0.094)	(0.086)	(0.079)	(0.059)	(0.079)	(0.061)	(0.077)	(0.064)	(0.086)
var	–0.214**	0.080	0.158	0.126	–0.059	–0.081	–0.096	–0.023	0.178	0.207
	(0.109)	(0.105)	(0.100)	(0.115)	(0.358)	(0.098)	(0.137)	(0.096)	(0.134)	(0.162)
Gas info x var	0.325***	–0.133	–0.067	–0.045	–0.044	0.052	0.103	0.164	0.001	–0.211*
	(0.121)	(0.121)	(0.118)	(0.118)	(0.464)	(0.108)	(0.151)	(0.107)	(0.128)	(0.112)
Gas info (var)	0.294***	0.026	0.052	0.063	0.043	0.081	0.144	0.135*	0.058	–0.028
	(0.096)	(0.075)	(0.080)	(0.088)	(0.461)	(0.077)	(0.140)	(0.078)	(0.113)	(0.073)
Enumerator FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Background variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	995	995	995	995	995	1989	1989	1989	1989	1989
R^2	0.206	0.202	0.201	0.201	0.201	0.175	0.176	0.176	0.175	0.177

Robust standard errors are in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: "In the years to come, I expect the extent of corrupt activities to: decrease a lot - decrease - stay the same - increase - increase a lot", standardization of five-point scale. *Treatment variable, interaction & background variables:* "Gas info" is an indicator variable taking a value of one for respondents who were exposed to gas information, "Gas info x var" is an interaction between "Gas info" and "Above median age" (H. age), "Lower sec or more" (H. educ), "Male", "Self-employed" (Self-empl) and "Farmer" respectively, "Gas info (var)" is the sum of estimated parameters for "Gas info" and the four respective indicator variables, "Male" is an indicator variable taking a value of one if the respondent is a male, "Above median age" is an indicator variable taking a value of one for respondents who are older than 34, "Lower sec or more" is an indicator taking a value of one if the respondent has completed lower secondary school or a higher level of education, "Self-employed" is a dummy taking a value of one for respondents who are self-employed in the non-agricultural sector (the reference category is a formal sector employee), "Farmer" is an indicator variable taking a value of one for respondents who are farmers (the reference category is a formal sector employee), "Married" is an indicator taking a value of one if the respondent is married and zero otherwise and "Gas region" is an indicator variable taking a value of one for respondents residing in Mtwara and Lindi (the reference category is Dar es Salaam). *Enumerator fixed effects (FE):* indicator variables for each of the 10 enumerators.

Table 6: Heterogeneity in the effect of gas information on novelty, for non-gas and gas regions

	Non-gas region					Gas regions				
	(1) H. age	(2) H. educ	(3) Male	(4) Self-empl	(5) Farmer	(6) H. age	(7) H. educ	(8) Male	(9) Self-empl	(10) Farmer
Gas info	0.162**	0.128	0.333***	0.194**	0.211***	–0.112	–0.170***	–0.169**	–0.046	–0.157*
	(0.068)	(0.084)	(0.079)	(0.082)	(0.057)	(0.082)	(0.064)	(0.085)	(0.068)	(0.084)
var	–0.288**	–0.587***	0.116	0.189*	–0.448	–0.030	–0.704***	–0.205*	0.460***	0.258*
	(0.113)	(0.100)	(0.100)	(0.113)	(0.397)	(0.106)	(0.134)	(0.105)	(0.151)	(0.154)
Gas info x var	0.171	0.175	–0.224**	0.060	0.984**	–0.001	0.356**	0.107	–0.312**	0.074
	(0.120)	(0.113)	(0.113)	(0.114)	(0.422)	(0.115)	(0.144)	(0.114)	(0.124)	(0.115)
Gas info (var)	0.333***	0.303***	0.109	0.254***	1.195***	–0.113	0.186	–0.062	–0.358***	–0.083
	(0.099)	(0.075)	(0.080)	(0.078)	(0.417)	(0.082)	(0.130)	(0.079)	(0.103)	(0.079)
Enumerator FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Background variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	995	995	995	995	995	1991	1991	1991	1991	1991
R^2	0.361	0.361	0.362	0.359	0.362	0.142	0.145	0.143	0.145	0.143

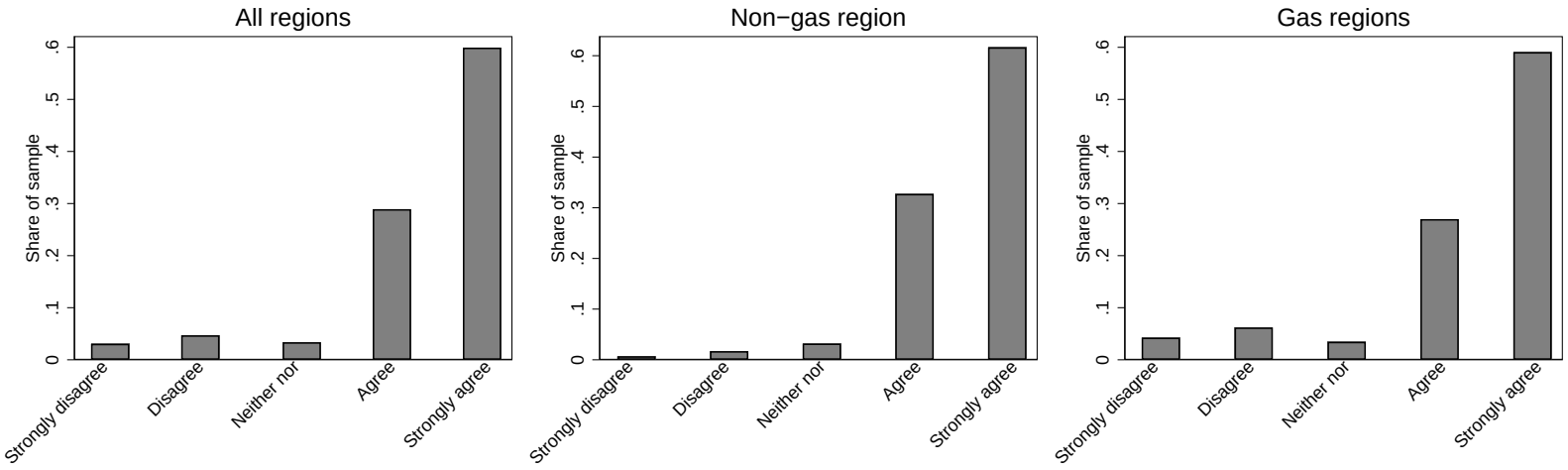
Robust standard errors are in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: "How much of this information was new to you? None of it - almost none of it - some of it - almost all of it - all of it", standardized five-point scale. *Treatment variable, interaction & background variables:* "Gas info" is an indicator variable taking a value of one for respondents who were exposed to gas information, "Gas info x var" is an interaction between "Gas info" and "Above median age" (H. age), "Lower sec or more" (H. educ), "Male", "Self-employed" (Self-empl) and "Farmer", respectively, "Gas info (var)" is the sum of estimated parameters for "Gas info" and the four respective indicator variables, "Male" is an indicator variable taking a value of one if the respondent is a male, "Above median age" is an indicator variable taking a value of one for respondents who are older than 34, "Higher education" is an indicator taking a value of one if the respondent has completed lower secondary school or a higher level of education, "Self-employed" is a dummy taking a value of one for respondents who are self-employed in the non-agricultural sector (the reference category is a formal sector employee), "Farmer" is an indicator variable taking a value of one for respondents who are farmers (the reference category is a formal sector employee), "Married" is an indicator taking a value of one if the respondent is married and zero otherwise and "Gas region" is an indicator variable taking a value of one for respondents residing in Mtwara and Lindi (the reference category is Dar es Salaam). *Enumerator fixed effects (FE):* indicator variables for each of the 10 enumerators.

Appendix A Additional figures and tables

Figure A.1: Selected regions

Figure A.2: Perceived corruption of others, by total, non-gas region and gas regions



Note: The figure illustrates the distribution of responses to the question "I believe that my fellow citizens would engage in corrupt activities if they could benefit from it: strongly disagree - agree - neither nor - agree - strongly agree", for the total sample (left panel), the non-gas region (middle panel) and the gas regions (right panel).

Table A.1: Balance regressions

	(1) All	(2) Non-gas region	(3) Gas regions
Above median age	0.004	0.023	–0.005
	(0.016)	(0.034)	(0.018)
Higher education	–0.003	–0.011	0.006
	(0.022)	(0.033)	(0.030)
Male	0.003	0.021	–0.002
	(0.015)	(0.030)	(0.017)
Self-employed	–0.014	0.002	–0.049
	(0.034)	(0.042)	(0.057)
Farmer	0.033	–0.019	0.013
	(0.037)	(0.133)	(0.054)
Married	0.025	0.048	0.013
	(0.017)	(0.034)	(0.020)
Rural region	–0.025		
	(0.039)		
Enumerator FE	Yes	Yes	Yes
Observations	2990	996	1994
R^2	0.054	0.007	0.064

Robust standard errors are in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Notes: Regressions performed with "Gas info", an indicator variable taking a value of one for individuals who were exposed to the gas information version of the video, as the dependent variable. *Background variables:* "Male" is an indicator variable taking a value of one if the respondent is a male, "Above median age" is an indicator variable taking a value of one for respondents who are older than 34, "Higher education" is an indicator taking a value of one if the respondent has completed lower secondary school or a higher level of education, "Self-employed" is a dummy taking a value of one for respondents who are self-employed in the non-agricultural sector (the reference category is a formal sector employee), "Farmer" is an indicator variable taking a value of one for respondents who are farmers (the reference category is a formal sector employee), "Married" is an indicator variable taking a value of one if the respondent is married and zero otherwise and "Gas region" is an indicator variable taking a value of one for respondents residing in Mtwara and Lindi (the reference category is Dar es Salaam). *Enumerator fixed effects (FE):* indicator variables for each of the 10 enumerators.

Table A.2: Observations by treatment

Treatment	Frequency	Per cent
No gas information	687	22.87
Yearly returns/per capita	714	23.77
Yearly return/population	671	22.34
Total value/per capita	519	17.28
Total value/population	413	13.75
Total	3004	100

Notes: The table illustrates the number of observations by treatment. The total/per capita and total value/population treatments have fewer observations because of challenges in undertaking the randomization using the app.

Table A.3: Effect on expected corruption of providing information with separate regressions for Mtwara and Lindi

	(1) All	(2) All	(3) Non-gas region	(4) Gas regions	(5) Mtwara	(6) Lindi
Gas info	0.077*	0.070*	0.086	0.057	–0.012	0.098
	(0.044)	(0.041)	(0.059)	(0.056)	(0.086)	(0.074)
Above median age		–0.024	0.020	–0.039	–0.061	–0.031
		(0.037)	(0.065)	(0.045)	(0.064)	(0.064)
Higher education		–0.021	–0.013	–0.016	0.015	–0.035
		(0.046)	(0.061)	(0.071)	(0.108)	(0.097)
Male		0.113***	0.111**	0.110***	0.052	0.165***
		(0.033)	(0.056)	(0.043)	(0.061)	(0.061)
Self-employed		0.128*	0.094	0.176	0.131	0.213
		(0.067)	(0.078)	(0.133)	(0.189)	(0.184)
Farmer		0.013	–0.090	0.036	0.012	0.062
		(0.077)	(0.247)	(0.129)	(0.188)	(0.178)
Married		–0.079**	–0.076	–0.082*	–0.045	–0.087
		(0.038)	(0.068)	(0.048)	(0.071)	(0.067)
Rural region		–0.016				
		(0.075)				
Enumerator FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2998	2984	995	1989	995	994
R^2	0.010	0.180	0.201	0.175	0.184	0.180

Robust standard errors are in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: "In the years to come, I expect the extent of corrupt activities to: decrease a lot - decrease - stay the same - increase - increase a lot", standardization of five-point scale. *Treatment variable, interaction & background variables:* "Gas info" is an indicator variable taking a value of one for respondents who were exposed to gas information, "Gas info x var" is an interaction between "Gas info" and "Above median age" (H. age), "Lower sec or more" (H. educ), "Male", "Self-employed" (Self-empl) and "Farmer", respectively, "Gas info (var)" is the sum of estimated parameters for "Gas info" and the four respective indicator variables, "Male" is an indicator variable taking a value of one if the respondent is a male, "Above median age" is an indicator variable taking a value of one for respondents who are older than 34, "Lower sec or more" is an indicator taking a value of one if the respondent has completed lower secondary school or a higher level of education, "Self-employed" is a dummy taking a value of one for respondents who are self-employed in the non-agricultural sector (the reference category is a formal sector employee), "Farmer" is an indicator variable taking a value of one for respondents who are farmers (the reference category is a formal sector employee), "Married" is an indicator taking a value of one if the respondent is married and zero otherwise and "Gas region" is an indicator variable taking a value of one for respondents residing in Mtwara and Lindi (the reference category is Dar es Salaam). *Enumerator fixed effects (FE):* indicator variables for each of the 10 enumerators.

Table A.4: Effect on corruption norm of providing information with separate regressions for Mtwara and Lindi

	Survey					**Behavior**				
	(1) All	(2) Non-gas region	(3) Gas regions	(4) Mtwara	(5) Lindi	(6) All	(7) Non-gas region	(8) Gas regions	(9) Mtwara	(10) Lindi
Gas info	–0.019	0.016	–0.043	–0.074	–0.022	–0.033	–0.049	–0.016	–0.136	0.052
	(0.037)	(0.045)	(0.053)	(0.091)	(0.062)	(0.048)	(0.070)	(0.063)	(0.099)	(0.083)
Above median age	0.051	0.109**	0.032	0.029	0.030	–0.588***	–0.431***	–0.670***	–0.626***	–0.715***
	(0.034)	(0.052)	(0.043)	(0.064)	(0.058)	(0.045)	(0.076)	(0.055)	(0.080)	(0.078)
Lower sec or more	0.061	0.053	0.057	0.135	–0.038	0.011	0.028	–0.002	0.070	–0.033
	(0.041)	(0.048)	(0.070)	(0.102)	(0.093)	(0.056)	(0.069)	(0.089)	(0.129)	(0.121)
Male	–0.022	–0.024	–0.024	–0.052	0.019	0.480***	0.476***	0.497***	0.513***	0.470***
	(0.031)	(0.042)	(0.042)	(0.062)	(0.056)	(0.043)	(0.072)	(0.055)	(0.078)	(0.078)
Self-employed	–0.085	–0.069	–0.038	–0.127	0.054	0.081	0.079	0.126	0.355*	–0.014
	(0.053)	(0.054)	(0.118)	(0.185)	(0.148)	(0.082)	(0.095)	(0.147)	(0.184)	(0.197)
Farmer	–0.189***	–0.492	–0.137	–0.117	–0.127	0.005	0.150	0.033	0.209	–0.216
	(0.065)	(0.301)	(0.115)	(0.181)	(0.147)	(0.095)	(0.231)	(0.142)	(0.186)	(0.191)
Married	0.091**	0.005	0.112**	0.126	0.114*	0.020	0.063	–0.040	–0.061	–0.024
	(0.036)	(0.049)	(0.049)	(0.079)	(0.061)	(0.047)	(0.079)	(0.060)	(0.092)	(0.083)
Rural region	0.122**					–0.140*				
	(0.061)					(0.082)				
Observations	2984	994	1990	996	994	1469	479	990	503	487
R^2	0.287	0.231	0.321	0.324	0.332	0.371	0.296	0.305	0.313	0.341

Robust standard errors are in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variables: Survey = "People should never engage in corrupt activities: strongly disagree - agree - neither nor - agree - strongly agree", standardization of five-point scale, and behavior = share of tails reported in the coin flipping game (0, 1/6, 2/6, 3/6, 4/6, 5/6, or 1), standardized values. *Treatment variable & background variables:* "Gas info" is an indicator variable taking a value of one for respondents who were exposed to gas information, "Male" is an indicator variable taking a value of one if the respondent is a male, "Above median age" is an indicator variable taking a value of one for respondents who are older than 34, "Higher education" is an indicator taking a value of one if the respondent has completed lower secondary school or a higher level of education, "Self-employed" is a dummy taking a value of one for respondents who are self-employed in the non-agricultural sector (the reference category is a formal sector employee), "Farmer" is an indicator variable taking a value of one for respondents who are farmers (the reference category is a formal sector employee), "Married" is an indicator taking a value of one if the respondent is married and zero otherwise and "Gas region" is an indicator variable taking a value of one for respondents residing in Mtwara and Lindi (the reference category is Dar es Salaam). *Enumerator fixed effects (FE):* indicator variables for each of the 10 enumerators.

Table A.5: Effect of providing information on perceived novelty of information, with separate regressions for Mtwara and Lindi

	(1) All	(2) All	(3) Non-gas region	(4) Gas regions	(5) Mtwara	(6) Lindi
Gas info	0.049	0.028	0.224***	–0.113*	–0.119	–0.091
	(0.046)	(0.042)	(0.056)	(0.058)	(0.092)	(0.075)
Above median age		–0.069*	–0.165***	–0.030	–0.018	–0.036
		(0.037)	(0.061)	(0.045)	(0.065)	(0.062)
Higher education		–0.437***	–0.465***	–0.430***	–0.313***	–0.527***
		(0.043)	(0.055)	(0.069)	(0.113)	(0.088)
Male		–0.094***	–0.040	–0.119***	–0.143**	–0.092
		(0.034)	(0.051)	(0.042)	(0.063)	(0.058)
Self-employed		0.226***	0.231***	0.236**	0.261	0.246
		(0.063)	(0.072)	(0.119)	(0.174)	(0.171)
Farmer		0.276***	0.235	0.318***	0.355**	0.307*
		(0.075)	(0.217)	(0.120)	(0.176)	(0.172)
Married		–0.027	–0.028	–0.031	–0.044	–0.031
		(0.038)	(0.057)	(0.049)	(0.075)	(0.066)
Gas regions		0.096				
		(0.068)				
Enumerator FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3000	2986	995	1991	995	996
R^2	0.014	0.189	0.359	0.142	0.132	0.171

Robust standard errors are in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: "How much of this information was new to you? None of it - almost none of it - some of it - almost all of it - all of it", standardized five-point scale. *Treatment variable, interaction & background variables:* "Gas info" is an indicator variable taking a value of one for respondents who were exposed to gas information, "Gas info x var" is an interaction between "Gas info" and "Above median age" (H. age), "Lower sec or more" (H. educ), "Male", "Self-employed" (Self-empl) and "Farmer", respectively, "Gas info (var)" is the sum of estimated parameters for "Gas info" and the four respective indicator variables, "Male" is an indicator variable taking a value of one if the respondent is a male, "Above median age" is an indicator variable taking a value of one for respondents who are older than 34, "Lower sec or more" is an indicator taking a value of one if the respondent has completed lower secondary school or a higher level of education, "Self-employed" is a dummy taking a value of one for respondents who are self-employed in the non-agricultural sector (the reference category is a formal sector employee), "Farmer" is an indicator variable taking a value of one for respondents who are farmers (the reference category is a formal sector employee), "Married" is an indicator taking a value of one if the respondent is married and zero otherwise and "Gas region" is an indicator variable taking a value of one for respondents residing in Mtwara and Lindi (the reference category is Dar es Salaam). *Enumerator fixed effects (FE):* indicator variables for each of the 10 enumerators.

Table A.6: Heterogeneity in effect of gas info on corruption norm, survey measure, for non-gas and gas regions

	Non-gas region					Gas regions				
	(1) Age	(2) H. educ	(3) Male	(4) Self-empl	(5) Farmer	(6) Age	(7) H. educ	(8) Male	(9) Self-empl	(10) Farmer
Gas info	–0.059	0.071	0.060	–0.069	0.008	–0.027	–0.057	–0.026	–0.012	–0.077
	(0.056)	(0.069)	(0.060)	(0.062)	(0.045)	(0.077)	(0.058)	(0.071)	(0.062)	(0.068)
var	–0.044	0.123	0.038	–0.191**	–0.933*	0.057	–0.006	0.002	0.065	–0.183
	(0.095)	(0.080)	(0.077)	(0.087)	(0.481)	(0.091)	(0.070)	(0.090)	(0.140)	(0.143)
Gas info x var	0.211**	–0.100	–0.089	0.173*	0.635	–0.031	0.082	–0.032	–0.144	0.056
	(0.098)	(0.091)	(0.090)	(0.090)	(0.588)	(0.100)	(0.115)	(0.100)	(0.108)	(0.100)
Gas info (var)	0.152*	–0.029	–0.029	0.105	0.644	–0.058	0.026	–0.059	–0.156*	–0.020
	(0.079)	(0.060)	(0.067)	(0.065)	(0.586)	(0.069)	(0.123)	(0.075)	(0.090)	(0.075)
Enumerator FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Background variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	994	994	994	994	994	1990	1990	1990	1990	1990
R^2	0.235	0.232	0.232	0.234	0.233	0.321	0.321	0.321	0.321	0.321

Robust standard errors are in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: "People should never engage in corrupt activities: strongly disagree - agree - neither nor - agree - strongly agree", standardization of five-point scale. *Treatment variable, interaction & background variables:* "Gas info" is an indicator variable taking a value of one for respondents who were exposed to gas information, "Gas info x var" is an interaction between "Gas info" and "Above median age" (H. age), "Lower sec or more" (H. educ), "Male", "Self-employed" (Self-empl) and "Farmer", respectively, "Gas info (var)" is the sum of estimated parameters for "Gas info" and the four respective indicator variables, "Male" is an indicator variable taking a value of one if the respondent is a male, "Above median age" is an indicator variable taking a value of one for respondents who are older than 34, "Lower sec or more" is an indicator taking a value of one if the respondent has completed lower secondary school or a higher level of education, "Self-employed" is a dummy taking a value of one for respondents who are self-employed in the non-agricultural sector (the reference category is a formal sector employee), "Farmer" is an indicator variable taking a value of one for respondents who are farmers (the reference category is a formal sector employee), "Married" is an indicator variable taking a value of one if the respondent is married and zero otherwise and "Gas region" is an indicator variable taking a value of one for respondents residing in Mtwara and Lindi (the reference category is Dar es Salaam). *Enumerator fixed effects (FE):* indicator variables for each of the 10 enumerators.

Table A.7: Heterogeneity in the effect of gas information on corruption norm, behavioral measure, for non-gas and gas regions

	Non-gas region					Gas regions				
	(1) Age	(2) H. educ	(3) Male	(4) Self-empl	(5) Farmer	(6) Age	(7) H. educ	(8) Male	(9) Self-empl	(10) Farmer
Gas info	–0.055	–0.041	–0.112	–0.008	–0.043	–0.022	0.002	–0.029	–0.037	0.012
	(0.089)	(0.116)	(0.113)	(0.099)	(0.071)	(0.086)	(0.070)	(0.095)	(0.073)	(0.092)
var	–0.441***	0.040	0.394***	0.138	0.357	–0.681***	0.079	0.478***	0.056	0.071
	(0.125)	(0.129)	(0.125)	(0.137)	(0.463)	(0.105)	(0.144)	(0.107)	(0.176)	(0.173)
Gas info x var	0.014	–0.016	0.111	–0.080	–0.306	0.013	–0.105	0.024	0.099	–0.047
	(0.146)	(0.144)	(0.142)	(0.140)	(0.507)	(0.122)	(0.167)	(0.125)	(0.141)	(0.125)
Gas info (var)	–0.041	–0.057	–0.001	–0.087	–0.350	–0.009	–0.103	–0.004	0.062	–0.035
	(0.114)	(0.084)	(0.087)	(0.098)	(0.501)	(0.089)	(0.152)	(0.083)	(0.121)	(0.086)
Enumerator FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Background variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	479	479	479	479	479	990	990	990	990	990
R^2	0.296	0.296	0.297	0.296	0.296	0.305	0.305	0.305	0.305	0.305

Robust standard errors are in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: Share of tails reported in the coin flipping game, standardized measure. *Treatment variable, interaction & background variables:* "Gas info" is an indicator variable taking a value of one for respondents who were exposed to gas information, "Gas info x var" is an interaction between "Gas info" and "Above median age" (H. age), "Lower sec or more" (H. educ), "Male", "Self-employed" (Self-empl) and "Farmer", respectively, "Gas info (var)" is the sum of estimated parameters for "Gas info" and the four respective indicator variables, "Male" is an indicator variable taking a value of one if the respondent is a male, "Above median age" is an indicator variable taking a value of one for respondents who are older than 34, "Higher education" is an indicator taking a value of one if the respondent has completed lower secondary school or a higher level of education, "Self-employed" is a dummy taking a value of one for respondents who are self-employed in the non-agricultural sector (the reference category is a formal sector employee), "Farmer" is an indicator variable taking a value of one for respondents who are farmers (the reference category is a formal sector employee), "Married" is an indicator taking a value of one if the respondent is married and zero otherwise and "Gas region" is an indicator variable taking a value of one for respondents residing in Mtwara and Lindi (the reference category is Dar es Salaam). *Enumerator fixed effects (FE):* indicator variables for each of the 10 enumerators.

Table A.8: Effect on expected corruption of providing information, robustness check

	(1) All	(2) All	(3) Non-gas regions	(4) Gas regions
Yearly returns/per capita	0.097*	0.097*	0.105	0.092
	(0.055)	(0.051)	(0.082)	(0.066)
Yearly return/population	0.026	0.018	0.067	–0.009
	(0.056)	(0.051)	(0.080)	(0.068)
Total value/per capita	0.035	0.034	0.024	0.030
	(0.058)	(0.053)	(0.084)	(0.070)
Total value/population	0.166***	0.142**	0.140*	0.139*
	(0.062)	(0.057)	(0.083)	(0.079)
Above median age		–0.024	0.018	–0.038
		(0.037)	(0.065)	(0.045)
Lower sec or more		–0.024	–0.017	–0.016
		(0.046)	(0.061)	(0.071)
Male		0.114***	0.107*	0.114***
		(0.033)	(0.056)	(0.043)
Self-employed		0.129*	0.096	0.178
		(0.067)	(0.078)	(0.134)
Farmer		0.013	–0.086	0.036
		(0.077)	(0.247)	(0.131)
Married		–0.081**	–0.081	–0.083*
		(0.038)	(0.068)	(0.048)
Rural region		–0.011		
		(0.076)		
p-value, F-test	0.101	0.105	0.630	0.161
Observations	2998	2984	995	1989
R^2	0.012	0.182	0.202	0.177

Robust standard errors are in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: "In the years to come, I expect the extent of corrupt activities to: decrease a lot - decrease - stay the same - increase - increase a lot", standardization of five-point scale. *Treatment variables:* "Annual returns/per capita" is an indicator variable taking a value of one for respondents exposed to the annual returns/per capita formulation, "Annual returns/population" is an indicator variable taking a value of one for respondents exposed to the annual returns/population formulation, "Total value/per capita" is an indicator variable taking a value of one for respondents exposed to the total value/per capita formulation, "Total value/population" is an indicator variable taking a value of one for respondents exposed to the total value/population formulation.*Background variables:* "Male" is an indicator variable taking a value of one if the respondent is a male, "Above median age" is an indicator variable taking a value of one for respondents who are older than 34, "Lower sec or more" is an indicator taking a value of one if the respondent has completed lower secondary school or a higher level of education, "Self-employed" is a dummy taking a value of one for respondents who are self-employed in the non-agricultural sector (the reference category is a formal sector employee), "Farmer" is an indicator variable taking a value of one for respondents who are farmers (the reference category is a formal sector employee), "Married" is an indicator taking a value of one if the respondent is married and zero otherwise and "Gas regions" is an indicator variable taking a value of one for respondents residing in Mtwara and Lindi (the reference category is Dar es Salaam). *Enumerator fixed effects (FE):* indicator variables for each of the 10 enumerators.

Table A.9: Effect on corruption norm of providing information, robustness check

	Survey			**Behavior**		
	(1) All	(2) Non-gas region	(3) Gas region	(4) All	(5) Non-gas region	(6) Gas region
Yearly returns/per capita	–0.047	–0.015	–0.075	–0.004	0.025	–0.007
	(0.047)	(0.065)	(0.063)	(0.062)	(0.096)	(0.079)
Yearly return/population	–0.018	0.002	–0.037	–0.086	–0.054	–0.087
	(0.047)	(0.059)	(0.066)	(0.063)	(0.092)	(0.081)
Total value/per capita	–0.003	0.039	–0.025	–0.058	–0.093	–0.033
	(0.050)	(0.066)	(0.068)	(0.066)	(0.104)	(0.086)
Total value/population	0.002	0.050	–0.020	0.027	–0.094	0.109
	(0.049)	(0.059)	(0.070)	(0.067)	(0.091)	(0.092)
Above median age	0.053	0.112**	0.033	–0.589***	–0.435***	–0.670***
	(0.034)	(0.052)	(0.043)	(0.045)	(0.076)	(0.055)
Lower sec or more	0.061	0.053	0.056	0.008	0.022	–0.004
	(0.041)	(0.048)	(0.070)	(0.056)	(0.068)	(0.089)
Male	–0.022	–0.024	–0.024	0.480***	0.476***	0.502***
	(0.031)	(0.042)	(0.042)	(0.043)	(0.072)	(0.055)
Self-employed	–0.087	–0.072	–0.039	0.081	0.081	0.123
	(0.053)	(0.055)	(0.119)	(0.083)	(0.096)	(0.147)
Farmer	–0.191***	–0.500*	–0.139	0.002	0.159	0.025
	(0.065)	(0.303)	(0.116)	(0.095)	(0.225)	(0.142)
Married	0.091**	0.004	0.111**	0.018	0.059	–0.044
	(0.036)	(0.049)	(0.049)	(0.047)	(0.078)	(0.061)
Rural region	0.122**			–0.135*		
	(0.061)			(0.082)		
p-value, F-test	0.777	0.756	0.815	0.359	0.697	0.221
Observations	2984	994	1990	1469	479	990
R^2	0.288	0.232	0.321	0.373	0.298	0.308

Robust standard errors are in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variables: Survey = "People should never engage in corrupt activities: strongly disagree - agree - neither nor - agree - strongly agree", standardization of five-point scale, and behavior = share of tails reported in the coin flipping game (0, 1/6, 2/6, 3/6, 4/6, 5/6, or 1), standardized values. *Treatment variables:* "Annual returns/per capita" is an indicator variable taking a value of one for respondents exposed to the annual returns/per capita formulation, "Annual returns/population" is an indicator variable taking a value of one for respondents exposed to the annual returns/population formulation, "Total value/per capita" is an indicator variable taking a value of one for respondents exposed to the total value/per capita formulation, "Total value/population" is an indicator variable taking a value of one for respondents exposed to the total value/population formulation.*Background variables:* "Male" is an indicator variable taking a value of one if the respondent is a male, "Above median age" is an indicator variable taking a value of one for respondents who are older than 34, "Lower sec or more" is an indicator taking a value of one if the respondent has completed lower secondary school or a higher level of education, "Self-employed" is a dummy taking a value of one for respondents who are self-employed in the non-agricultural sector (the reference category is a formal sector employee), "Farmer" is an indicator variable taking a value of one for respondents who are farmers (the reference category is a formal sector employee), "Married" is an indicator taking a value of one if the respondent is married and zero otherwise and "Gas regions" is an indicator variable taking a value of one for respondents residing in Mtwara and Lindi (the reference category is Dar es Salaam). *Enumerator fixed effects (FE):* indicator variables for each of the 10 enumerators.

Table A.10: Effect on perceived novelty of providing information, robustness check

	(1) All	(2) All	(3) Non-gas region	(4) Gas regions
Yearly returns/per capita	0.049	0.050	0.257***	–0.094
	(0.056)	(0.051)	(0.076)	(0.068)
Yearly return/population	0.128**	0.073	0.260***	–0.064
	(0.056)	(0.052)	(0.071)	(0.070)
Total value/per capita	0.011	–0.020	0.181**	–0.163**
	(0.061)	(0.055)	(0.080)	(0.072)
Total value/population	–0.022	–0.013	0.178**	–0.156**
	(0.064)	(0.056)	(0.076)	(0.079)
Above median age		–0.071*	–0.169***	–0.032
		(0.037)	(0.061)	(0.045)
Lower sec or more		–0.435***	–0.465***	–0.426***
		(0.043)	(0.055)	(0.069)
Male		–0.095***	–0.038	–0.121***
		(0.034)	(0.052)	(0.042)
Self-employed		0.231***	0.236***	0.240**
		(0.063)	(0.073)	(0.119)
Farmer		0.283***	0.246	0.326***
		(0.075)	(0.217)	(0.120)
Married		–0.026	–0.026	–0.029
		(0.038)	(0.058)	(0.049)
Rural region		0.095		
		(0.068)		
Enumerator FE	Yes	Yes	Yes	Yes
p-value, F-test	0.068	0.233	0.626	0.390
Observations	3000	2986	995	1991
R^2	0.016	0.190	0.360	0.144

Robust standard errors are in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: How much of this information was new to you? None of it - almost none of it - some of it - almost all of it - all of it", standardized five-point scale. *Treatment variables:* "Annual returns/per capita" is an indicator variable taking a value of one for respondents exposed to the annual returns/per capita formulation, "Annual returns/population" is an indicator variable taking a value of one for respondents exposed to the annual returns/population formulation, "Total value/per capita" is an indicator variable taking a value of one for respondents exposed to the total value/per capita formulation, "Total value/population" is an indicator variable taking a value of one for respondents exposed to the total value/population formulation. *Background variables:* "Male" is an indicator variable taking a value of one if the respondent is a male, "Above median age" is an indicator variable taking a value of one for respondents who are older than 34, "Lower sec or more" is an indicator taking a value of one if the respondent has completed lower secondary school or a higher level of education, "Self-employed" is a dummy taking a value of one for respondents who are self-employed in the non-agricultural sector (the reference category is a formal sector employee), "Farmer" is an indicator variable taking a value of one for respondents who are farmers (the reference category is a formal sector employee), "Married" is an indicator taking a value of one if the respondent is married and zero otherwise and "Gas regions" is an indicator variable taking a value of one for respondents residing in Mtwara and Lindi (the reference category is Dar es Salaam). *Enumerator fixed effects (FE):* indicator variables for each of the 10 enumerators.

Table A.11: Effect on corruption beliefs of providing information

	(1) All	(2) Non-gas regions	(3) Gas regions
Gas info	0.008	0.007	0.000
	(0.034)	(0.052)	(0.044)
Above median age	–0.081***	–0.002	–0.113***
	(0.029)	(0.061)	(0.034)
Lower sec or more	0.005	–0.007	–0.011
	(0.039)	(0.056)	(0.057)
Male	–0.035	0.110**	–0.101***
	(0.027)	(0.049)	(0.033)
Self-employed	0.055	0.064	–0.022
	(0.056)	(0.064)	(0.114)
Farmer	0.057	–0.047	0.021
	(0.062)	(0.140)	(0.108)
Married	–0.040	–0.020	–0.040
	(0.031)	(0.059)	(0.038)
Rural region	0.062		
	(0.054)		
Observations	2981	993	1988
R^2	0.477	0.514	0.473

Robust standard errors are in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: "I believe that my fellow citizens would engage in corrupt activities if they could benefit from it: strongly disagree - agree - neither nor - agree - strongly agree", standardization of five-point scale. *Treatment variable & background variables:* "Gas info" is an indicator variable taking a value of one for respondents who were exposed to gas information, "Male" is an indicator variable taking a value of one if the respondent is a male, "Above median age" is an indicator variable taking a value of one for respondents who are older than 34, "Higher education" is an indicator taking a value of one if the respondent has completed lower secondary school or a higher level of education, "Self-employed" is a dummy taking a value of one for respondents who are self-employed in the non-agricultural sector (the reference category is a formal sector employee), "Farmer" is an indicator variable taking a value of one for respondents who are farmers (the reference category is a formal sector employee), "Married" is an indicator taking a value of one if the respondent is married and zero otherwise and "Gas region" is an indicator variable taking a value of one for respondents residing in Mtwara and Lindi (the reference category is Dar es Salaam). *Enumerator fixed effects(FE):* indicator variables for each of the 10 enumerators.

Table A.12: Heterogeneity in effect of gas information on corruption beliefs, for non-gas and gas regions

	Non-gas region					Gas regions				
	(1) H. age	(2) H. educ	(3) Male	(4) Self-empl	(5) Farmer	(6) H. age	(7) H. educ	(8) Male	(9) Self-empl	(10) Farmer
Gas info	0.036	–0.014	0.095	–0.000	0.001	0.015	–0.011	–0.125**	–0.031	0.133**
	(0.067)	(0.078)	(0.069)	(0.073)	(0.052)	(0.064)	(0.048)	(0.062)	(0.051)	(0.063)
var	0.056	–0.035	0.236***	0.054	–0.378*	–0.091	–0.063	–0.294***	–0.128	0.202
	(0.101)	(0.094)	(0.089)	(0.096)	(0.198)	(0.076)	(0.103)	(0.075)	(0.128)	(0.127)
Gas info x var	–0.081	0.039	–0.181*	0.015	0.478**	–0.027	0.068	0.239***	0.146	–0.224***
	(0.105)	(0.104)	(0.104)	(0.103)	(0.232)	(0.084)	(0.117)	(0.083)	(0.098)	(0.085)
Gas info (var)	–0.044	0.025	–0.086	0.015	0.479**	–0.012	0.057	0.114*	0.115	–0.091
	(0.081)	(0.069)	(0.077)	(0.073)	(0.226)	(0.059)	(0.108)	(0.059)	(0.084)	(0.059)
Enumerator FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Background variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	993	993	993	993	993	1988	1988	1988	1988	1988
R^2	0.514	0.514	0.515	0.514	0.514	0.473	0.473	0.475	0.473	0.475

Robust standard errors are in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent variable: "I believe that my fellow citizens would engage in corrupt activities if they could benefit economically from it: strongly disagree - agree - neither nor - agree - strongly agree", standardization of five-point scale. *Treatment variable, interaction & background variables:* "Gas info" is an indicator variable taking a value of one for respondents who were exposed to gas information, "Gas info x var" is an interaction between "Gas info" and "Above median age" (H. age), "Lower sec or more" (H. educ), "Male", "Self-employed" (Self-empl) and "Farmer", respectively, "Gas info (var)" is the sum of estimated parameters for "Gas info" and the four respective indicator variables, "Male" is an indicator variable taking a value of one if the respondent is a male, "Above median age" is an indicator variable taking a value of one for respondents who are older than 34, "Higher education" is an indicator variable taking a value of one for respondents who have completed lower secondary schooling or more, "Self-employed" is a dummy taking a value of one for respondents who are self-employed in the non-agricultural sector (the reference category is a formal sector employee), "Farmer" is an indicator variable taking a value of one for respondents who are farmers (the reference category is a formal sector employee), "Married" is an indicator taking a value of one if the respondent is married and zero otherwise and "Gas region" is an indicator variable taking a value of one for respondents residing in Mtwara and Lindi (the reference category is Dar es Salaam). *Enumerator fixed effects (FE):* indicator variables for each of the 10 enumerators.

Appendix B Supplementary materials

B.1 Survey

Good day. My name is NN and I am from REPOA, an independent research organisation based in Dar es Salaam. I do not represent the government or any political party. We are studying the views of citizens in Tanzania. We would like to discuss these issues with a member of your household. The information obtained here will be treated strictly confidentially. The answers to these questions will be an important input when it comes to prescribing policies to improve the system.

First, we want to know a little bit about you.

Q1.3 What is your marital status?

- Married
- Widow/Widower
- Unmarried
- Divorced

Q1.4 How old are you?

Q1.5. What is your "home" region/region of origin?

Q1.6 To which region do you go most frequently to visit relatives?

Q1.7 What is the highest level of schooling you have completed?

- No formal schooling
- Some primary school
- Primary school completed
- Lower secondary (form 1-4)
- College (after lower secondary)
- Upper secondary (form 5-6)
- College (after upper secondary)
- Vocational/adult education classes
- Some university or university completed
- Don't know

Q1.8 What is your main occupation?

- Central government employee
- Local government employee
- Private company employee, international company
- Private company employee, domestic company
- Self-employed (non-agriculture)
- Employed in NGO/CSO
- Unemployed
- Farmer
- Student
- Retired
- Don't know
- Other (specify in English) _____

[App randomizes respondent to gas info or non-gas info version of video]

I will now show you a video with some information about Tanzania on my tablet. The video has both pictures and sound. When the video starts playing, please let me know if you have problems seeing the pictures or hear the sound. [Hand tablet and headset to respondent and help with adjustment of headset]. Press play when you are ready. Please pay careful attention.

Thank you watching the video. Let us now continue with the rest of the questions.

Q2.3 Did you find the estimated gas revenues larger than expected? (Only gas information group)

- Yes
- No

Q7.1 How much of this information was new to you?

- All of it
- Almost all of it
- Some of it
- Almost none of it
- None of it

Q8.1 For each of the following statements, please complete / whether you agree or disagree.

Q8.2 In order for Tanzania to achieve social and economic development in the years to come, the taxes paid by ordinary Tanzanians should

- Decrease a lot
- Decrease
- Stay the same
- Increase
- Increase a lot

Q8.3 If a political party advocates an increase in taxes paid by ordinary Tanzanian citizens my support for that party will

- Decrease a lot
- Decrease
- Stay the same
- Increase
- Increase a lot

Q8.4 In the years to come, I expect the provision of public services, such as schooling, health services and roads to

- Worsen a lot
- Worsen
- Stay the same
- Improve
- Improve a lot

Q8.5 Difference in income between rich and poor in Tanzania should

- Decrease a lot
- Decrease
- Stay the same
- Increase
- Increase a lot

Q8.6 If a political party advocates a reduction in differences in income between rich and poor my support for that party will

- Decrease a lot
- Decrease
- Stay the same
- Increase
- Increase a lot

Q8.7 In the years to come, I expect differences in income between rich and poor to

- Decrease a lot
- Decrease
- Stay the same
- Increase
- Increase a lot

Q8.8 In the years to come, I trust the government to do what is right for Tanzania

- Strongly disagree
- Disagree
- Neither Agree nor Disagree
- Agree
- Strongly Agree

Q8.9 Generally, I trust my fellow citizens

- Strongly disagree
- Disagree
- Neither Agree nor Disagree
- Agree
- Strongly Agree

Q8.10 The well-being of my fellow citizens is important to me

- Strongly disagree
- Disagree
- Neither Agree or Disagree
- Agree
- Strongly agree

Q8.11 The well-beeing of future generations is important to me

- Strongly disagree
- Disagree
- Neither Agree nor Disagree
- Agree
- Strongly Agree

Q8.12 People should never engage in corrupt activities

- Strongly disagree
- Disagree
- Neither Agree nor Disagree
- Agree
- Strongly Agree

Q8.13 I believe that my fellow citizens would engage in corrupt activities if they could benefit economically from it

- Strongly disagree
- Disagree
- Neither Agree nor Disagree
- Agree
- Strongly Agree

Q8.14 In the years to come, I expect the extent of corrupt activities to

- Decrease a lot
- Decrease
- Stay the same
- Increase
- Increase a lot

[App randomizes respondent to play trust OR cheating game]

Trust game

As part of this research project, we also study how Tanzanians make economic choices. We will therefore ask you to make an economic choice that has real consequences. In this part of the study, you will be paired with a randomly selected Tanzanian citizen. We have already asked this Tanzanian citizen what he or she would do in the situation we are now going to present to you and we have pre-registered his or her response. As we told this person, we will use his or her response to determine the payment to a number of the people we interview, but in the following instructions we simplify the presentation and only describe how this will work for you.

The money paid to you in this part of the study is determined by your decsions and comes without any obligations or conditions. The sole purpose of the payments is to study economic decisions.

I will now give you Tsh 5000. [Give the respondent Tsh 5000 in Tsh 500 notes]. You can choose between keeping the entire amount to yourself, sending the whole amount to the random Tanzanian citizen, or spilt the amount between the two of you in any way that you want (in portions of Tsh 500). If you choose to send any money to the Tanzanian citizen, the amount you send will be tripled in size, meaning that for every Tsh 1000 you send, your Tanzanian citizen will receive Tsh 3000.

Just to make sure that you have understood the task, I will ask you two questions.

Q58 If you send Tsh 2000 to the randomly selected Tanzanian, how much will he or she receive?

[If respondent answers Tsh 6000, please proceed to the next question].

If respondent answers anything else than Tsh 6000, please say the following:

Your answer is not correct. I will explain one more time. The amount you send will be tripled in size, meaning that for every Tsh 1000 you send to your Tanzanian citizen, he or she will receive Tsh 3000. I will ask the question once again. If you send Tsh 2000 to the randomly selected Tanzanian you have been paired with, how much will he or she receive? Repeat this until the respondent answer correctly, Tsh 6000. Then proceed to the second question]

That is correct. Let me ask you the question in a different way.

For every Tsh 1000 you send to the randomly selected Tanzanian, how much will he or she receive?

[If the respondent answers Tsh 3000, please proceed to the allocation decision]

[If respondent answers anything else than Tsh 3000, say the following: Your answer is not correct. I will explain one more time. The amount you send will be tripled in size, meaning that for every Tsh 1000 you send, your Tanzanian citizen will receive Tsh 3000. I will ask the question once again. For every Tsh 1000 you send to the randomly selected Tanzanian you have been paired with, how much will he or she receive? Repeat this until the respondent answer correctly, Tsh 3000. Then proceed to the allocation decision]

Q60 I would now like you to tell me how much, if anything, you would like to send to your Tanzanian citizen. Please take this amount from the Tsh 5000 and give it to me.

[Respondent hands money over to you. Write down the amount given to you and check the payment table to determine the amount you should give to the respondent]

Thank you for your decision. The randomly selected Tanzanian citizen has told us that if receiving this amount of money, he or she returns Tsh XXX back to you.

[Hand Tsh XXX to the respondent and note down the amount in the box below. Amount sent and sent back should also be noted down on your separate payment sheet]

Cheating game

As part of this research project, we also study some economic situations that have real consequences. I now invite you to perform a task. First, I will demonstrate what I want you to do.

[Show respondent a Tsh 200 coin]

This coin has two sides. One with the picture of KARUME and one with the picture of a LION. I would like you to flip the coin like this

[Flip the coin, grab it and put it on top of your wrist. Holding the other hand on top of the coin. Then lift your upper hand to look at which side of the coin is up, but make sure that the coin is only visible to you, and not to the respondent] and to repeat this 6 times. I will ask you to do the flipping of the coin in privacy, and then come back and report the number of times the LION side was up. When you have completed the task, I will pay you Tsh 1000 for each LION.

The money paid to you in this part of the study is determined by the way you toss the coin and comes without any obligations or conditions. The sole purpose of the payments is to study economic situations.

Just to make sure that the task is understood, I will ask you two questions:

Q63 If you get LION 5 times, how much will I pay you?

[If respondent answers Tsh 5000, proceed to second question.]

[If respondent answers anything else than Tsh 5000, say the following:

Your answer is not correct. Let me explain again. For every LION you get, I will pay you 1000 Tsh. Let me ask the question again. If you get LION 5 times, how much will I pay you? Repeat this until the respondent answer correctly, Tsh 5000. Then proceed to the second question]

Q64 That is correct. Let me ask you in a different way. How much will I pay you for each LION?

[If respondent answers Tsh 1000, proceed to second demonstration.]

[If respondent answers anything else than Tsh 1000, say the following:
Your answer is not correct. Let me explain again. For every LION you get, I will pay you 1000]

[Proceed to second demonstration]

I will demonstrate once more before you can proceed to flip the coin.

[Repeat demonstration of coin flipping, and hand the coin over to the respondent]

Please flip the coin 6 times and count how many times the LIONS side is up. You will do the coin flips in privacy. Please step away while you complete the task and return to me when you are done.

[Wait for the respondent to complete and call him/her back]

Q66 Thank you for completing this task for me. How many LIONS did you get?

[Note down the number of LIONS in the box below. Also note the amount paid to the respondent on the separate payment sheet. Complete the survey by saying:]

All participants
This was the final part of the interview. Thank you for completing this task for me and for answering my questions. Your input is of great value to our research!

[Fill in the next section after the interview has been completed]

Q9.2 Gender of respondent

- Man
- Woman

Q9.3 District

1. Ilala (Dar es Salaam) (1)
2. Temeke (Dar es Salaam) (2)
3. Kiniodoni (Dar es Salaam) (3)
4. Lindi (Lindi) (4)
5. Lindi Manispaa (Lindi) (5)
6. Nachingwea (Lindi) (7)
7. Masasi Vijijini (Mtwara) (10)
8. Mtwara Manispaa (Mtwara) (12)
9. Newala (Mtwara) (15)

Q9.4 County/ward (specify)

Q9.5 Local / Village (specify)

Q10.1 Name of interviewer

Q10.2 Age of interviewer

Q10.3 Sex of interviewer

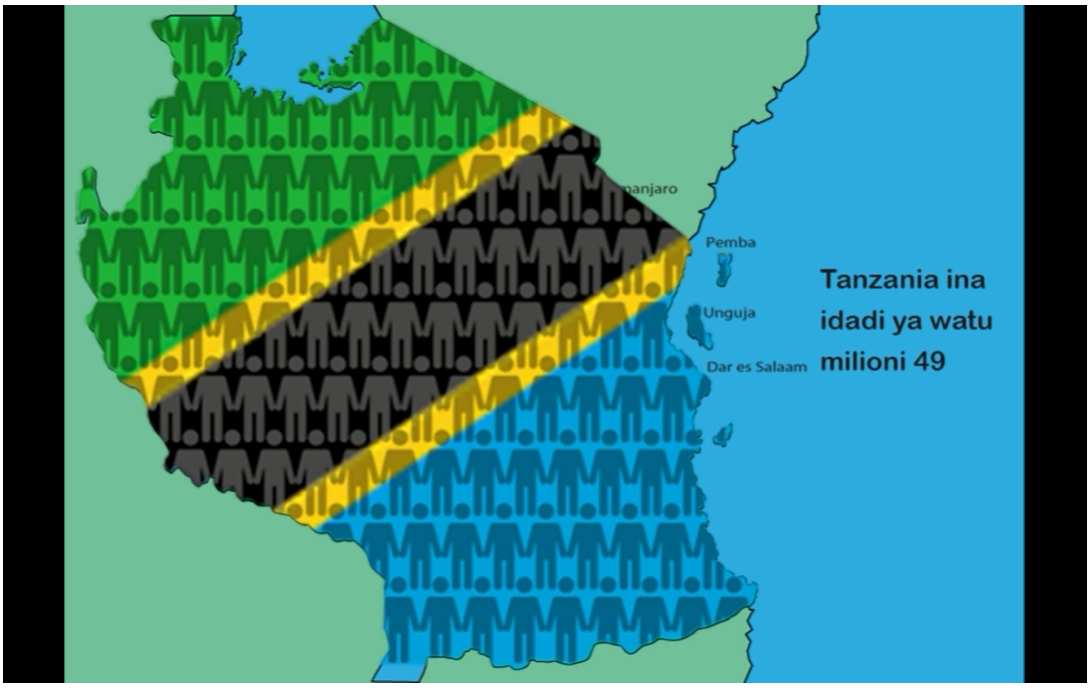
- Man
- Woman

B.2 Videos

Figure B.1: Slide 1 of video, general information (all respondents)

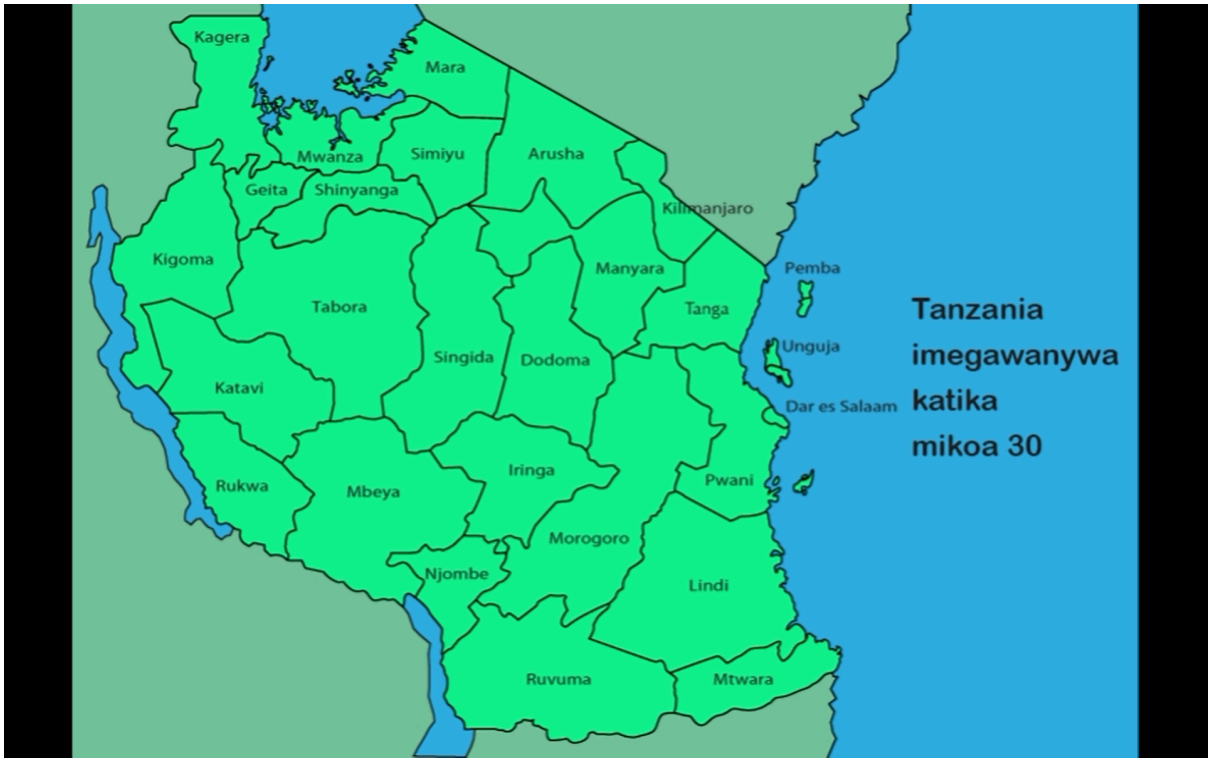
Note: The figure illustrates the first slide in the general information part of the video. The text (and sound) translates to: "This video will give you some information about Tanzania".

Figure B.2: Slide 2 of video, general information (all respondents)



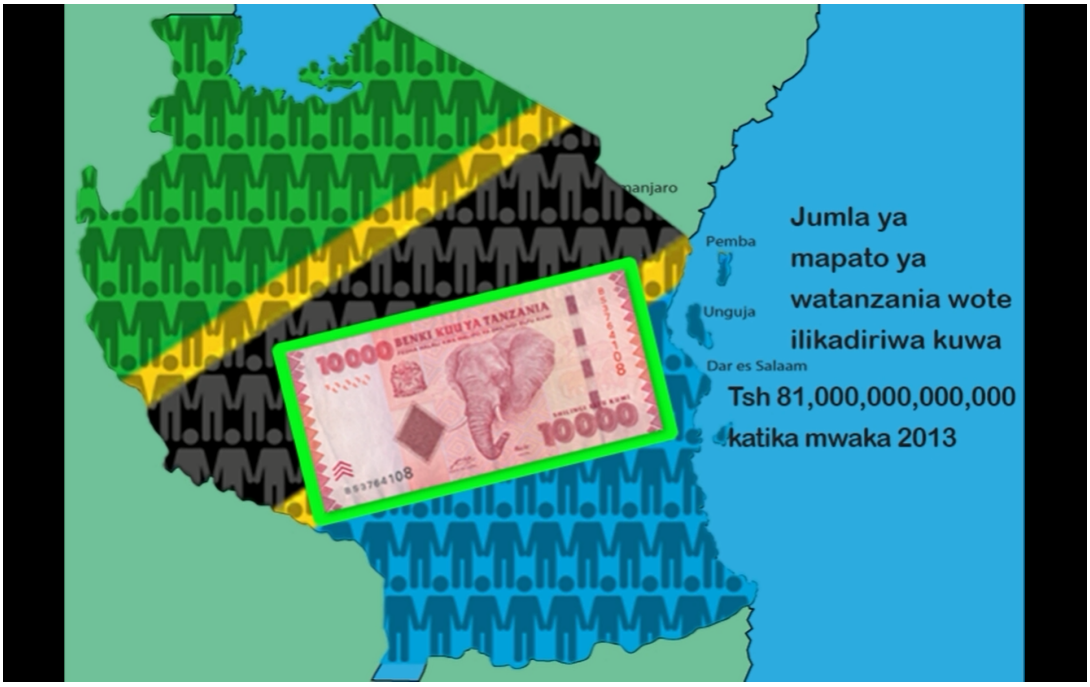
Note: The figure illustrates the second slide in the general information part of the video. The text (and sound) translates to: "Tanzania has a population of 49 million".

Figure B.3: Slide 3 of video, general information (all respondents)



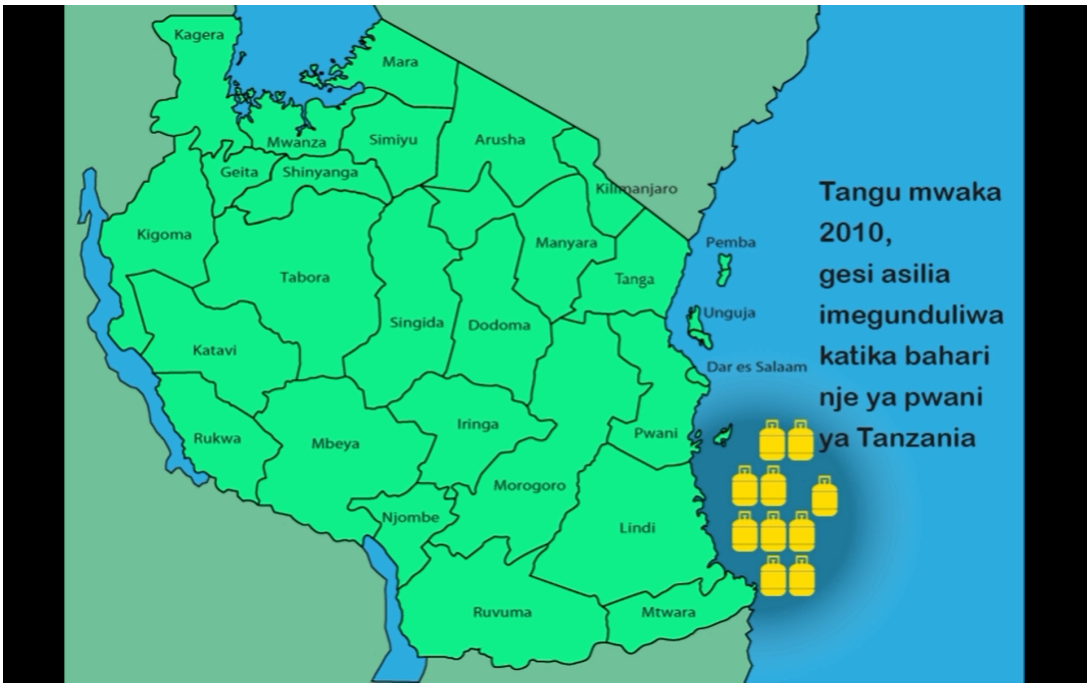
Note: The figure illustrates the third slide in the general information part of the video. The text (and sound) translates to: "Tanzania is divided into 30 regions".

Figure B.4: Slide 4 of video, general information (all respondents)



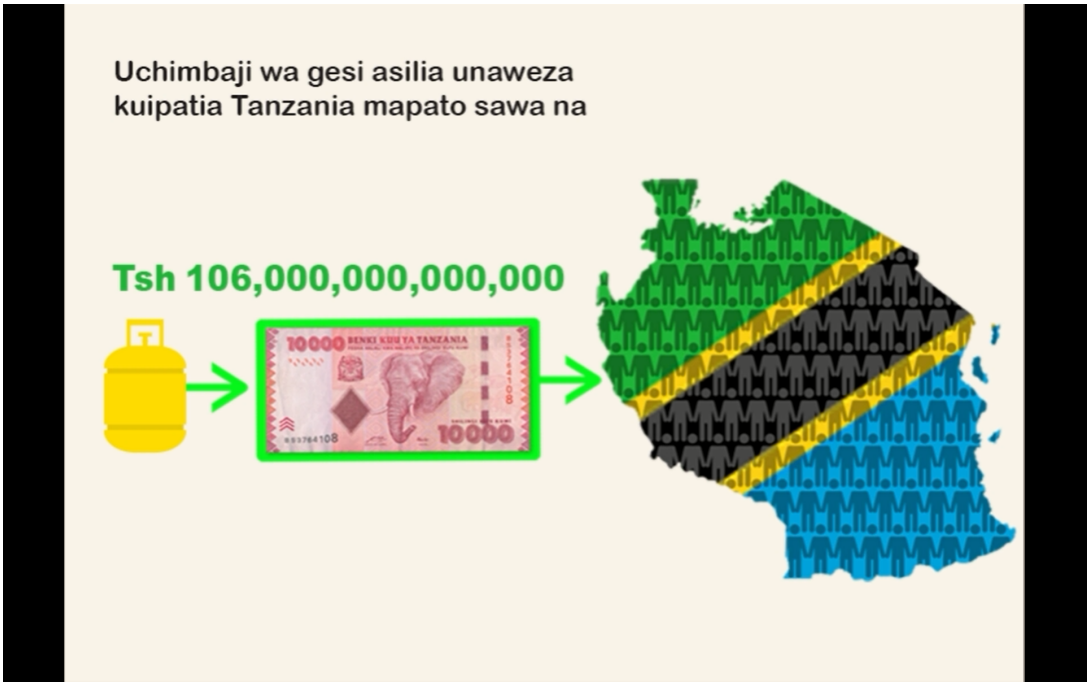
Note: The figure illustrates the fourth slide in the general information part of the video. The text (and sound) translates to: "Total revenue for all Tanzanians was estimated to be 81 trillion Tanzania shillings in 2013".

Figure B.5: Slide 5 of video, gas information (gas information group only, all formulations)



Note: The figure illustrates the first slide in the gas information part of the video. The text (and sound) translates to: "Since the year 2010, natural gas has been discovered in the ocean offshore of the Tanzania coast".

Figure B.6: Slide 6 of video, gas information (gas information group only, all formulations)

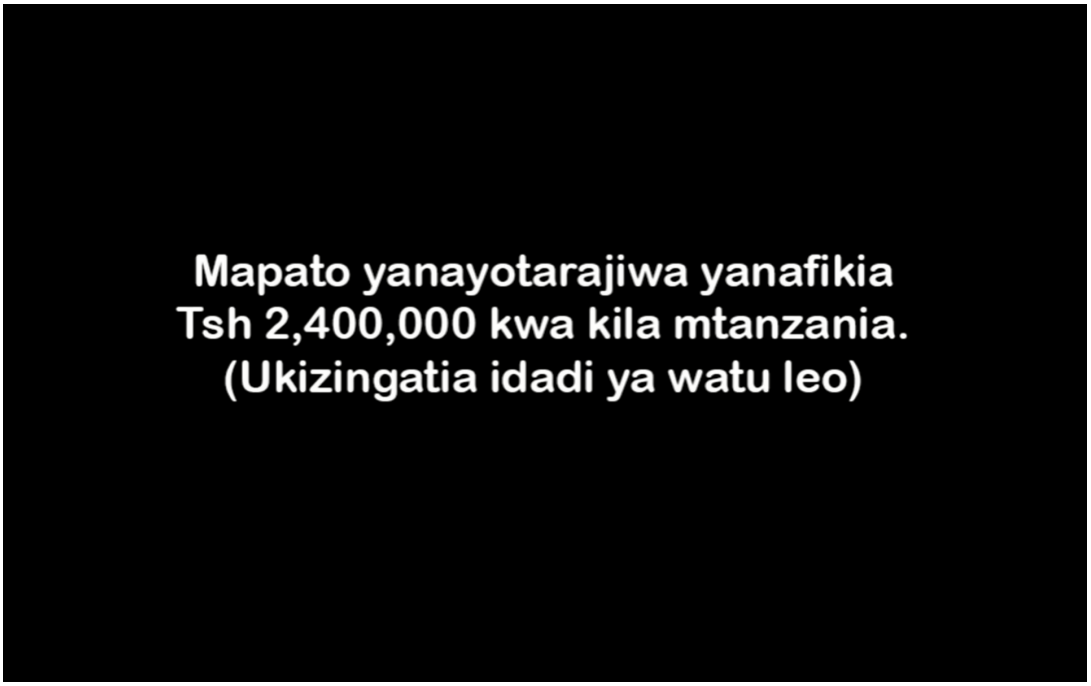


Note: The figure illustrates the second slide in the gas information part of the video. The text (and sound) translates to: "This gas can be sold and yield Tanzania a total revenue of TZS 106 trillion".

Figure B.7: Slide 7 of video, gas information (Total value/per capita, yearly return/per capita and yearly return/population formulations)

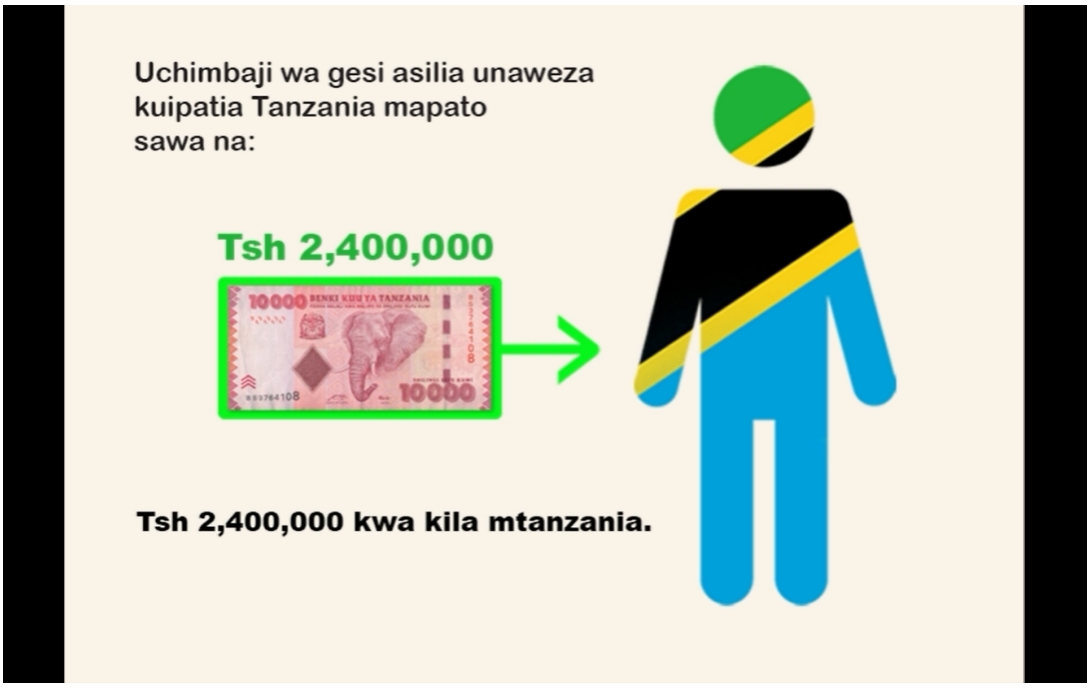
Note: The figure illustrates the second slide in the gas information part of the video. The text (and sound) translates to: "Now, let us look at this revenue in a different way".

Figure B.8: Slide 8 of video, gas information, total value/per capita formulation



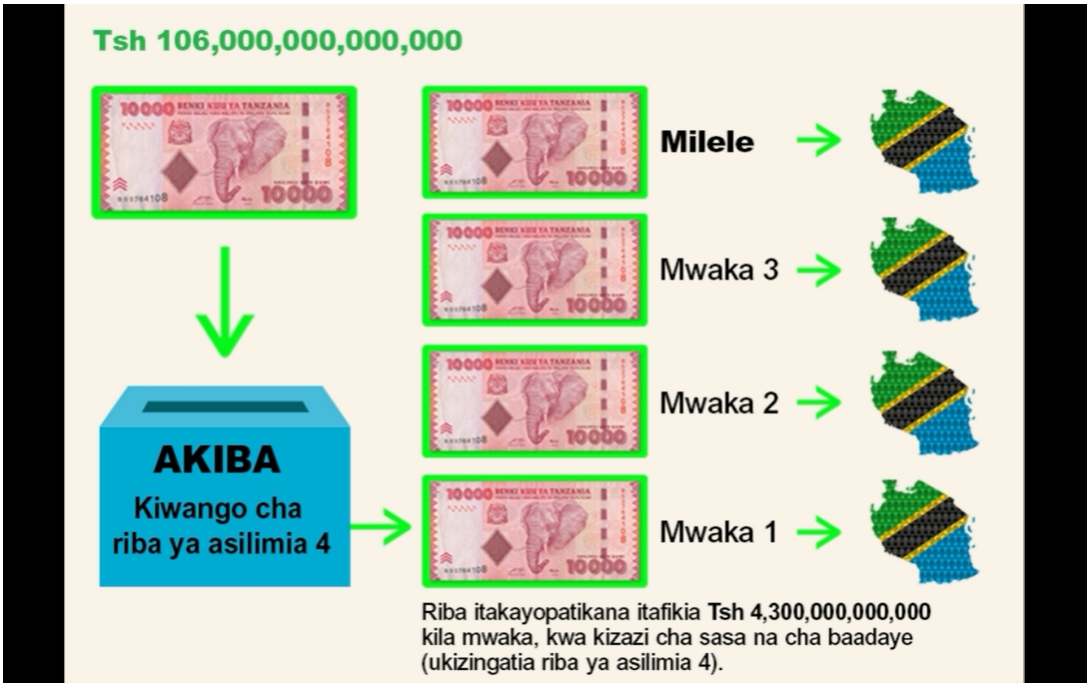
Note: The figure illustrates the first slide in the total value/per capita formulation version of the gas information. The text (and sound) translates to: "The expected revenue amounts to TZS 2.4 million per Tanzanian (considering the population today)".

Figure B.9: Slide 9 of video, gas information, total value/per capita formulation



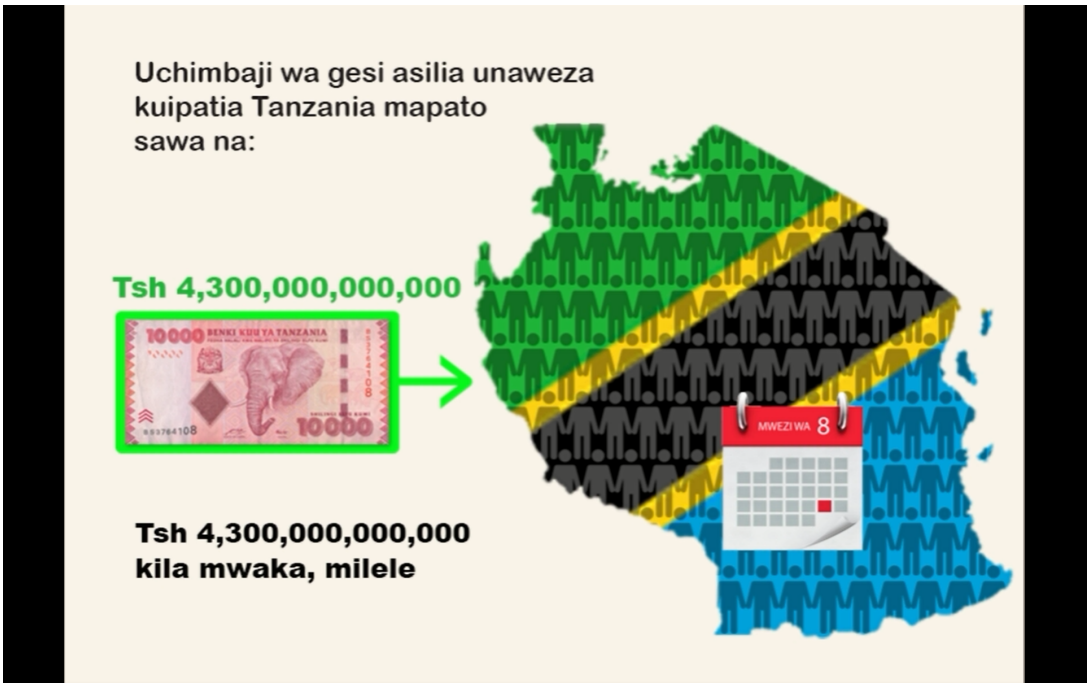
Note: The figure illustrates the second slide in the total value/per capita formulation version of the gas information. The text (and sound) translates to: "The extraction of natural gas can yield Tanzania a revenue equal to TZS 2.4 million to each Tanzanian".

Figure B.10: Slide 8 of video, gas information, yearly return/population formulation



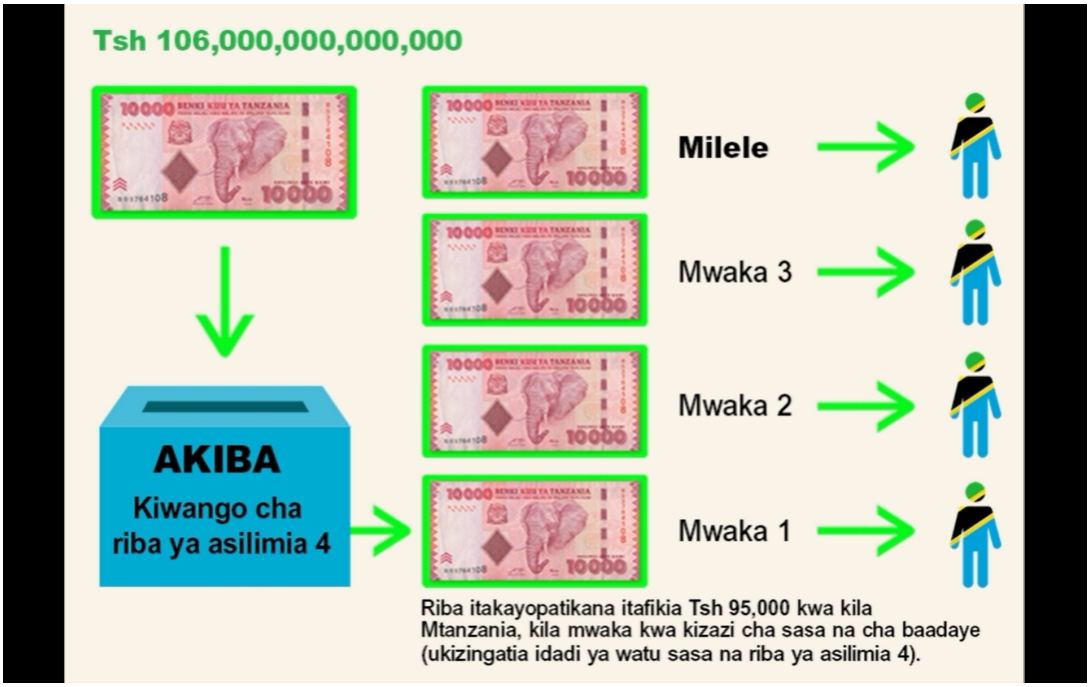
Note: The figure illustrates the first slide in the yearly return/population formulation version of the gas information. The text (and sound) translates to: "If we, as a nation, take the revenues expected from the sale of natural gas; and deposit them at an interest rate of 4%, the interest obtained will amount to TZS 4.3 trillion each year for the present and future generations".

Figure B.11: Slide 9 of video, gas information, yearly return/population formulation



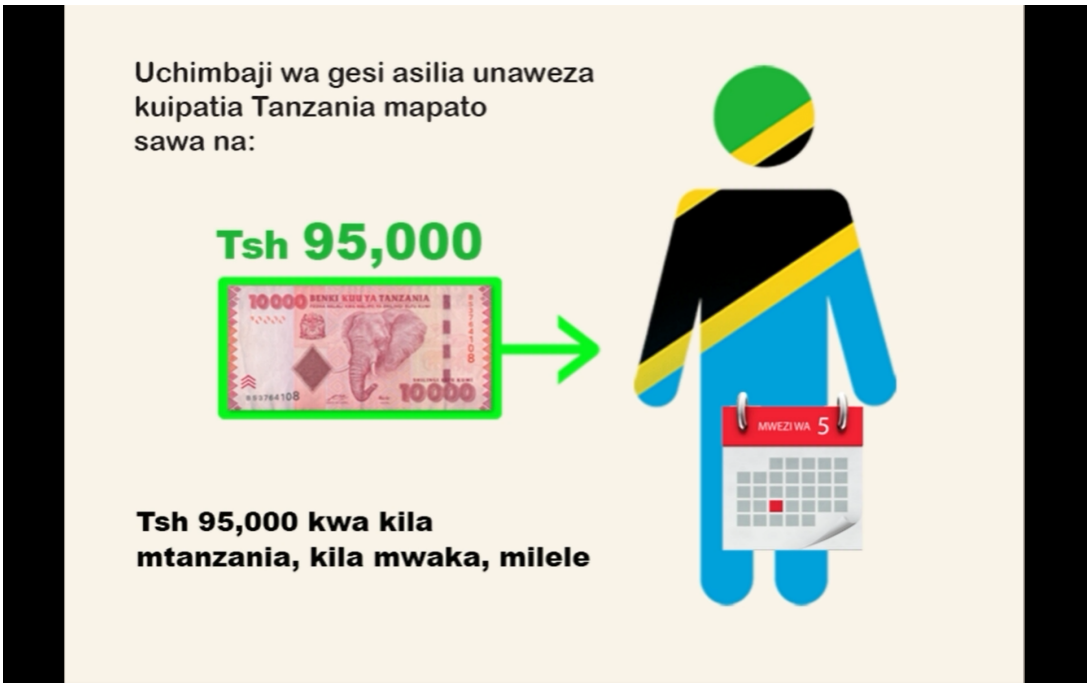
Note: The figure illustrates the second slide in the yearly return/population formulation version of the gas information. The text (and sound) translates to: "The extraction of natural gas can yield Tanzania revenue equal to TZS 4.3 trillion each year, forever".

Figure B.12: Slide 8 of video, gas information, yearly return/per capita formulation



Note: The figure illustrates the first slide in the yearly return/population formulation version of the gas information. The text (and sound) translates to: "If we take these revenues expected from the sale of natural gas; and deposit them at an interest rate of 4%, the interest obtained will amount to TZS 95,000 for each Tanzanian each year for the present and future generations, considering the present population".

Figure B.13: Slide 9 of video, gas information, yearly return/per capita formulation



Note: The figure illustrates the second slide in the yearly return/population formulation version of the gas information. The text (and sound) translates to: "The extraction of natural gas can give Tanzania the revenue equal to TZS 95,000 for each Tanzanian, each year for ever".

Figure B.14: Last slide of video (all respondents)

Note: The figure illustrates the last slides of all videos. The text (and sound) translates to: "Thank you for listening carefully, we would now like to ask you some questions" .

Corruption appears to be an important driver of the resource curse in developing countries. We report from a large-scale eld experiment in Tanzania that provides causal evidence on how expectations about future natural resource revenues shape expectations about corruption and the willingness to engage in corrupt behavior. We find robust evidence that information about the discovery of natural gas inTanzania causes people to expect more corruption in the future, but has no impact on people's willingness to engage in corrupt behavior. We believe that our results shed some light on underlying mechanisms of the resource curse.

This Working Paper is an output from *Tanzania as a future petrostate: Prospects and challenges*, a five-year (2014–19) institutional collaborative programme for research, capacity building, and policy dialogue. It is jointly implemented by REPOA and CMI, in collaboration with the National Bureau of Statistics. The programme is funded by the Norwegian Embassy, Dar es Salaam.

ISSN 0804-3639 (print) ISSN 1890-5048 (PDF) ISBN 978-82-8062-687-5 (print) ISBN978-82-8062-688-2 (PDF)