Why efforts to fight corruption can undermine the social contract

Cheeseman, Nic; Peiffer, Caryn

Link to publication on Research at Birmingham portal

Governance WILEY

Why efforts to fight corruption can undermine the social contract: Lessons from a survey experiment in Nigeria

Nic Cheeseman[1] | **Caryn Peiffer[2]**

[1]International Development Department, University of Birmingham, Birmingham, UK

[2]International Public Policy and Governance, School for Policy Studies, University of Bristol, Bristol, UK

Correspondence
Nic Cheeseman, International Development Department, University of Birmingham, Birmingham, UK.
Email: n.cheeseman@bham.ac.uk

Funding information
FCDO

Abstract

Anticorruption awareness raising efforts are designed to encourage citizens to resist and report corruption but have been found to either not work or have unwanted effects—including increasing bribe payment. This article represents the first test of whether these efforts also undermine critical aspects of a society's social contract, namely, willingness to pay tax. Using a household level survey experiment in Lagos, Nigeria, we assess whether exposure to five messages about (anti)corruption influence citizens' belief that they have a duty to pay taxes, or "tax morale". Though they were different in tone and content, four of the five messages undermined tax morale. We argue that this is likely because anti-corruption messages raise awareness of corruption risks, and hence concerns that taxes will be wasted. In turn, this highlights a new potential unintended and unwanted consequence of policy interventions that focus on raising awareness of government failings.

Two of the main concerns of those researching what are often call "developing countries" over the last 30 years have been high levels of corruption and the absence of a social contract between citizens and governments. The magnitude of corruption has been a constant preoccupation due to the potential for graft and financial mismanagement to undermine government revenues and public services (Igiebor, 2019; Lawal, 2007; Ogundiya, 2017). Similarly, the weakness—some

would argue absence (Prichard, 2015)—of a strong social contract in some countries has drawn attention because it is argued to be central to effective democratic government (Abbink, 2009; Prichard, 2019). Relatedly, a number of scholars have pointed to the limited direct income tax paid by large sections of society—especially in regions such as sub-Saharan Africa where the informal sector makes up most of the economy—as both a proxy of a social contract and a driver of poor governance (Bodea & LeBas, 2016; Moore et al., 2018; Prichard, 2019). On this argument, people who do not pay direct taxes are less motivated to make demands for good governance because it is not clear that money lost to waste or corruption was once their own (Prichard, 2015). In turn, if citizens do not expect—and indeed demand—a certain standard of performance from elected officials they will not hold them accountable, facilitating incompetence. The reverse is also true: Paler's (2013, p. 707) study of Indonesia found that raising participant's awareness that the government manages "their" taxes "increased monitoring and anti-incumbent political action".

Although these two issues are often studied separately, they are intimately connected. On the one hand, weak social contracts can encourage corruption by undermining accountability, giving political leaders little incentive to reform (Hoffman & Patel, 2017). Conversely, rampant corruption can undermine citizens' confidence in the state (Burbidge, 2016), encouraging individuals to opt out of nascent social contracts (McCullough et al., 2021). Taken together, these processes have the potential to create a vicious cycle that encourages corruption while undermining the kind of checks and balances required if graft is to be brought under control. Yet although these interconnected mechanisms mean that corruption can generate a range of social and political "bads" that extend far beyond economic inefficiency, little research has been conducted on the impact of anti-corruption campaigns on public attitudes toward citizenship and the duties that it implies. As Andersson and Heywood (2012, p. 48) argue, there is a pressing "need for a better understanding of the relationship between anti-corruption campaigns and governance, and their impact on democracy".

This need is particularly urgent given that existing research on anticorruption awareness raising, a prominent feature of most anticorruption strategies globally, finds that they often have unintended and unwanted effects—including increasing bribe payment (Corbacho et al., 2016; Peiffer, 2017, 2018; Peiffer and Walton, 2019). This article builds on this research and responds to Andersson and Heywood's call by providing the first ever test of whether anticorruption messages unintentionally undermine critical aspects of a society's social contract, namely, willingness to pay tax, or "tax morale" (Bodea & LeBas, 2016). As Ortega et al. (2016) have argued in the case of Latin America, tax payment is shaped not only by the coercive capacity of the government but also by how it is perceived. We investigate the relationship between anti-corruption messages and tax morale using data from a unique and purpose-designed household-level survey experiment conducted in Lagos, Nigeria. In the study, a 2400-person representative sample of Lagosians were randomly exposed to one of five messages about (anti)corruption based on the kinds of messages that are commonly used by civil society groups and international donors, or no message at all (the control group). Each of the five treatments was designed to have a very different and specific focus, variously emphasizing the fact that: (1) corruption is widespread (*widespread*); (2) religious leaders stand against corruption (*religious*); (3) the government has had success in fighting corruption (*government success*); (4) the local impact of corruption (*local impact*); and, (5) the impact of corruption on tax revenues (*taxes*). Following exposure to these messages (or not for the control group), all respondents were asked questions about their experiences of and attitudes toward corruption, and how willing they were to pay tax.

Drawing on theoretical insights from political psychology research regarding message priming, and existing studies of anti-corruption messages, we hypothesize that all messages—regardless of different tones or content—are likely to have the same effect: reduce tax morale. Political psychologists have found that messages can (unintentionally) cause people to think more about their pre-existing beliefs about an issue, especially for topics like corruption, for which pre-existing beliefs tend to be strongly held (Carmines & Stimson, 1989; Lenz, 2009). In the particular context of our study, this gives rise to the expectation that *regardless of their differences*, any message about corruption may prime widely held pre-existing notions that graft is systemic, and hence that personal tax payments will be wasted. In line with this, recent studies of anticorruption awareness raising campaigns have found that very different types of message have similar negative effects, for example, undermining citizens' sense that they can make a difference in fighting corruption and even encouraging bribery (Cheeseman & Peiffer, 2021; Peiffer, 2018). Drawing together these two literature, we hypothesize that anticorruption messaging will undermine adherence to well-known but costly aspects of the social contract, such as tax morale.

Our results confirm this expectation with one exception. Four of the five messages undermined a belief that citizens should pay taxes, and did so to a similar degree. Only the religious treatment did not impact on willingness to pay tax. Drawing on further survey questions to arbitrate between alternative explanations, we suggest that this is likely to be because the religious treatment is the only one that does not explicitly prime respondents to think about the state. While all of the other treatments explicitly reference government officials, public services, politicians and/or civil servants, the religious treatment solely focuses on the efforts of religious leaders to combat corruption and does not mention where corruption occurs. We therefore argue that this narrative encourages recipients to worry about corruption but not to lay the blame at the feet of state officials, with the consequences that unlike the other messages this treatment does not undermine tax morale.

While further research is necessary to confirm the generalizability of these findings, this study suggests that anti-corruption messages may have a wider set of negative externalities than was previously thought. It is not just bribery that can be encouraged by awareness raising campaigns, but also willingness to perform actions such as tax payment—one of the most critical duties that fall upon citizens. Our analysis also enables us to refine and nuance the conclusions of previous research in this area. All anti-corruption messages may not have an equal impact on all political attitudes; rather, our results suggest that messages that effectively connect corruption to the government are particularly harmful. Put another way, the negative impact of anti-corruption messages is likely strongest when two things are primed simultaneously: the idea that corruption is pervasive, and the notion that the state is the locus of corruption.

If this is the case, there is a way to minimize the negative externalities of anticorruption work, but it is one that leaves corruption campaigners in a bind. To do the least harm policy makers and activists should avoid explicitly invoking the government and the state in their campaigns. The obvious problem with this solution is that if anti-corruption organizations follow this advice their messages may appear to shield precisely those individuals and institutions at the heart of the problem. Yet without taking this kind of precaution, there is a serious risk that anti-corruption messaging will encourage citizens to stop investing in their government, intensifying the challenge of state building. While not paying taxes may be a rational behavior in a corrupt state, it may also harm citizen's interests in the long-term, because governments have been found to be more likely to invest tax revenues "toward expenditures that benefit citizens" than other forms of income (Gadenne, 2017, p. 274).

1 | THEORIZING TAXATION, CORRUPTION AND THE SOCIAL CONTRACT

There are several approaches to conceptualizing social contracts and understanding how they emerge. As Sam Hickey (2011, p. 428) has observed, these range from a more liberal approach drawing on Hobbes (1651) and Hayek (2014) that emphasizes rational individuals coming together and agreeing to sacrifice some of their liberties to maximize security and utility as a community, to the more rights-based approach stressed by Rousseau (2018) and Rawls (Hampton, 1980; Rawls, 1971), which emphasizes individuals coming together out of mutual respect to build a just and equal commonwealth. One thing that all strands agree on, however, is that reciprocity—that is, mutually beneficial social cooperation—is fundamental to the evolution of a successful state. In the context of Western European state-building (Tilly, 1990) social contracts are typically argued to have emerged through a process of negotiation between those seeking to govern and those they hope to govern over. As Bodea and LeBas summarize (2016, p. 3), this literature stresses "the importance of iterated bargaining between states and societal actors. The outcome of this bargaining is the establishment of a mutually beneficial social contract, in which citizens defer to the authority of the state, pay taxes and receive public goods".

Over the last 2 decades, there has been a growing consensus within the development literature that negotiated social contracts are necessary for sustainable and inclusive development. This trend is driven by "a set of normative and ideological concerns within international development and an increased awareness of the importance of politics in shaping development policy outcomes" (Hickey, 2011, p. 427). The same literature is also clear that corruption represents a significant challenge to the evolution of such a contract. Corruption suggests that citizens' scarce resources, in the form of tax revenue, are not being well used, undermining the principle of reciprocity. Yenni Mangoting et al. (2015, p. 967) argue that the creation of a social contract in Russia following the collapse of the Soviet Union, was undermined by "high and massive corruption and government failure to provide public services". Similarly, Masana Ndinga-Kanga et al. suggest that the weak social contract between citizens and the African National Congress (ANC) government in South Africa is rooted in "declining trust in government institutions to deliver, including because they are riddled with corruption" (Ndinga-Kanga et al., 2020, p. 33). Significantly, this relationship appears to hold in both democratic and non-democratic contexts. Shantayanan Devarajan and Elena Ianchovichina (2017) argue that growing corruption was one of the key factors that broke the social contract in a number of Arab countries in the 2000s, leading to what came to be known as the Arab Spring.

While citizen compliance in a range of areas is important to effective governance, tax payment is generally seen to be particularly significant to the social contract (Bodea & LeBas, 2016). The willingness to give up our hard-earned wages to the state in return for protection and public services is one of the toughest tests—perhaps second only to support for military conscription—of public commitment to the political system. Indeed, for some scholars, tax compliance is by its very nature "a social contract between ordinary citizens and the government" (Uslaner, 2010, p. 175).

It is therefore not surprising that research on the impact of corruption on tax compliance reflects the broader literature on corruption and the social contract in microcosm. Perhaps most significantly, there is a broad consensus that perceptions of widespread corruption undermine tax revenues by eroding "tax morale" (D'Arcy, 2011). Especially in contexts with a large informal sector in which the enforcement of taxation is weak, closing the "tax gap"—the difference between the total tax due to the government and what is actually paid—requires voluntary public

compliance (Prichard, 2015). In this context, evidence that state revenues are being diverted from their intended destination; used to fund the activities of the ruling party; and/or, that powerful individuals use their connections to escape payment, is particularly problematic because it is likely to undermine citizens' willingness to declare their own taxes. Evidence of corruption within the state has also been argued to signal to citizens that it is easy to get away with not paying taxes and that they have less of a duty to do so. Thus, Jahnke and Weisser (2019, p. 1) conclude, on the basis of data from sub-Saharan Africa, that "petty corruption not only has a direct effect on tax morale but also diminishes confidence in tax authorities and therefore affects tax morale indirectly".

The experience of Nigeria provides an empirical example of this dynamic. A history of government waste and corruption has undermined public confidence in key institutions and the willingness of individuals to pay tax (Meagher, 2018; Umar et al., 2018). According to a survey conducted by Afrobarometer in 2017, over three-quarters (77%) of Nigerians believed that it is likely that "a rich person could pay a bribe or use personal connections to get away with avoiding paying taxes they owe to government". Given this, it is unsurprising that only 24% of Nigerians strongly agreed that "the tax authorities always have the right to make people pay taxes" (Afrobarometer, 2018). In this way, higher levels of corruption are expected to be associated with lower public tax morale.

2 | POLITICAL PSYCHOLOGY, ANTI-CORRUPTION CAMPAIGNS, AND TAX MORALE

The negative association between corruption and tax morale in the literature suggests that anti-corruption messages highlighting the problem of corruption in government may undermine willingness to pay tax. But combining theoretical developments within political psychology with the latest anti-corruption research suggest a much starker hypothesis: that any message about corruption may undermine the social contract.

Where political psychology is concerned, recent experimental studies have established that messaging can shape attitudes and behavior through "priming" (e.g., Berinsky et al., 2011; Carter et al., 2011). A message "primes" an issue when it simply causes its audience think more about it. This is different to "persuasion" or "learning", which is when messaging causes people to think differently about a topic, for instance by giving them new information (Brody & Page, 1972; Lenz, 2009; Riker, 1986). Intuitively, researchers have found that messages are most likely to shape attitudes by priming when it comes to issues they already hold strong views about (Carmines & Stimson, 1989; Lenz, 2009). On such topics, an individual is less likely to update their point of view when presented with a message and more likely to instead to recall their strongly held pre-existing beliefs. This can be reinforced by a second mechanism, namely, that when someone feels strongly about an issue they tend to focus on the information they receive about that specific topic, rather than other information that may be provided in the same message (Lang, 2000; Pfau et al., 1997). In line with this, cognitive response research has even shown that a message's content can have no impact on whether pre-existing attitudes are recalled (Cacioppo & Petty, 1989; Eagly & Chaiken, 1993).

This is important, because it means that a message intended to change what people believe about an issue (e.g., by providing new information about how well the government has done in closing the tax gap) could unintentionally cause people to recall contrary and undesired beliefs (e.g., that corruption is embedded throughout government). This is particularly significant for

our purposes because corruption is precisely the kind of topic on which people already hold strong views. It is not only a common subject of conversation, but politicians around the globe invoke the term corruption when campaigning. The global push to fight corruption has meant that citizens in many so-called "developing" countries have been exposed to anticorruption programming for decades.

Sure enough, the literature shows that anti-corruption awareness raising campaigns likely generate unintended and unwanted outcomes. Although these efforts are motivated by the idea that raising awareness about the problem of corruption—via billboards, posters, murals, and so on—will encourage people to resist corruption (i.e., see United Nations, 2004), in many cases they have been found to have no effects and in some cases make the situation worse. To date, five main studies have been conducted that have systematically assess the impact of anti-corruption messages using experimental techniques. In the first, Corbacho et al. (2016) conducted a survey experiment in Costa Rica and found that those exposed to a message that described bribery increasing were more likely to indicate willingness to bribe on a survey than those who were not. In the second, Peiffer's (2017, 2018) survey experiment in Jakarta, Indonesia, found that exposure to four different messages—which highlighted the pervasiveness of corruption, government successes in fighting corruption, and how citizens could fight against corruption—*all* increased worry about corruption, decreased pride in the government, and reduced the belief that ordinary people could easily participate in anti-corruption movements (Peiffer, 2018). The third paper reached similar conclusions. Cheeseman and Peiffer (2020) employed a "bribery" game in which participants had an opportunity to win a small amount of money to test the impact of five messages in Nigeria. They found that none of the messages reduced willingness to bribe, and that in most cases exposure made participants more likely to pay a bribe.

Although the remaining two studies did not generate analogous negative findings, they also concluded that in general anti-corruption messages did not have a positive effect. Peiffer and Walton's (2019) survey experiment in Port Moresby, Papua New Guinea (PNG), found that three out of four messages had no impact on attitudes toward reporting corruption. Meanwhile, in the fifth article Kobis et al. (2019) deployed a lab-in-the-field experiment in Manguzi, KwaZulu-Natal, South Africa, with participants playing a game in which they took on the roles of a citizen or a public official. Posters with a message about how bribery had declined were then erected in the town part way through the study. Although Kobis et al. (2019) found that participants playing the role of the official became less likely to choose to bribe after the posters were displayed, there was no positive effect for participants who played the role of the citizen. Given that very few participants were officials in real life, whereas all were citizens, their overall finding was in line with Peiffer and Walton's (2019).

To summarize, none of the existing studies found an overall positive effect of anti-corruption messaging, and a majority concluded that exposure to messages about corruption can backfire by increasing willingness to tolerate and participate in corrupt behaviors, or decreasing willingness to report corruption. Significantly, these unintended impacts have been associated with exposure to both "negatively" framed messages, such as those that emphasize how widespread corruption is (Cheeseman & Peiffer, 2020; Corbacho et al., 2016; Peiffer, 2017, 2018), and "positively" framed messages, like those that emphasize how citizens can get involved in anticorruption activities (Cheeseman & Peiffer, 2020; Peiffer, 2017, 2018). In other words, the "backfiring" documented in these studies suggests that any message at all about corruption may have unintended effects on attitudes and behaviors toward corruption.

What does this mean for the as yet untested relationship between anti-corruption messages and the willingness of citizens to pay taxes? Bringing together the theoretical insights from

prior work on corruption and the social contract, research on priming from political psychology, and the pessimistic findings of research on anti-corruption messaging, we hypothesis that any anti-corruption message is likely to have a negative impact on tax morale. By raising awareness to the problem, messaging campaigns likely prime—or reinforce—notions of corruption being a widespread and intractable problem. In priming negative stereotypes of politicians and the "system", messaging campaigns may therefore inadvertently cause citizens to recall a belief that the taxes will be wasted. Who would want to pay tax if they think it will be stolen?

3 | METHODOLOGY

To test this proposition, we use data from an original experiment conducted in Lagos, Nigeria from December 21, 2019 to January 12, 2020. Lagos was selected for a number of reasons. First, corruption is a serious problem in Nigeria. Transparency International's (2019) Corruption Perception Index ranks Nigeria 146/183 for its control of corruption (Transparency International, 2020). Second, corruption is not a socially taboo topic to discuss in Nigeria (Smith, 2008), and so recruitment of potential participants was not problematic. Third, the Lagos State government has had some well publicized "wins" in its efforts to collect more taxes and use them to build infrastructure and provide services (Bodea & LeBas, 2016), most notably under the Governorship of Babatunde Fashola (2007–2015). Although some of these gains were subsequently reversed, and there have been accusations that Fashola misused public funds while in office (Akinloye, 2015), this means that it is feasible citizens will believe reform is possible—if not always realized. Lagos therefore represents a sterner test of our hypothesis than areas in which no progress has been made over the last 20 years, where citizens are likely to be more pessimistic governments use of tax revenues.

Finally, despite recent economic growth, it is estimated that about 5.5 million people, or around three-quarters, of the workforce operates in the informal sector. Along with many individuals who operate as private consultants and advisors, this creates a large number of citizens who do not pay income tax through automated PAYE deductions by their employer. Voluntary tax compliance for individuals at all levels of wealth therefore remains very much a government priority and a public issue, with official events held to praise those who comply and embarrass those who do not (Cheeseman & de Grammont, 2017). Lagos is also a diverse location in terms of wealth, ethnicity and religion, and so features many of the conditions that are present in large cities in many developing countries. This means that our findings are likely to have resonance beyond this specific case—although we recognize that it is important to be cautious when generalizing on the basis of one study.

3.1 | Design

Working with Practical Sampling International (PSI), an experienced research firm in Lagos, we adopted Afrobarometer's established recruitment protocol, where random sampling was employed at every stage (Afrobarometer, 2020) to recruit a 2572-participant representative-of-Lagos sample for this study. Details on the sampling strategy and demographic characteristics of the sample are available in Appendix A.

Participants were randomly assigned to one of six groups: *control, widespread corruption, religious, government success, local fight* or *taxes* (n = 421–434 in each). For each interview,

professional enumerators from PSI started by reading a short introduction that explained that the study wanted to "learn what citizens think about public services and the experiences they have with public officials". Participants were then told they could withdraw at any time and responses would be confidential.

All participants were first asked the same demographic questions. If assigned to the *widespread corruption, religious, government success, local fight* or *taxes* groups, respondents were then asked to read their group's respective treatment paragraph (message). Following exposure to the treatment (or not for the control group, which proceeded immediately), participants were asked a series of survey questions to gauge their perceptions of corruption levels, anti-corruption efforts, willingness to engage in anti-corruption activism, political attitudes, and attitudes toward paying taxes, which we focus on in this article.

3.2 | Treatments

Although we hypothesize that anti-corruption messages have a common effect, we test five treatments to be as comprehensive and systematic as possible in light of the fact that this is the first experimental study to look at the relationship between such messages and tax morale. As noted above, in contrast to past studies in related areas, which leverage small variations between messages to reveal which terms or micro-signals are most/least effective, our expectation was that any message on corruption would have the same effect. We therefore designed messages that were substantially different from one another. The messages were selected either because they reflect current practice in anti-corruption campaigns (widespread; religious; government success), or because past research suggests they may have distinctive effects (local fight; taxes). Each of the treatments were a paragraph long (see Appendix B for wording).[1]

The *widespread* message highlights the pervasive nature of corruption in Nigeria. This treatment is thus designed to raise awareness to the extent of the issue, in line with the advice of awareness raising advocates (United Nations, 2004). Meanwhile, the *religious* treatment describes the outcry against corruption from different Nigerian faith leaders and is used because some anti-corruption campaigners have called for religious figures to be prominent in the fight against corruption on the basis that such "messengers" will be especially persuasive in areas of widespread religiosity (Marquette, 2012).

The *government success* treatment emphasizes some of the achievements claimed by the Lagos State Government in fighting and reducing corruption. It was tested because a current theme in anti-corruption efforts suggests that people may feel especially pessimistic about the chances of corruption reducing if they do not believe their government can control it, and so messaging that highlights improvements may encourage others to mobilize. For its part, the *local fight* treatment frames corruption as an issue that impacts local communities and resources, stating that local, communal efforts should be prioritized first. It was inspired by Peiffer and Walton's (2019) study in PNG, which found that a similar message encouraged more positive attitudes toward reporting corruption. As in PNG, many people identify strongly with their local and/or ethnic communities in Nigeria.

Finally, the *taxes* treatment argues that corruption represents the theft of taxes and fees that ordinary citizens pay daily. As this message explicitly highlights the connection between corruption and taxation, one might expect that exposure to it will be especially impactful on attitudes about tax payment. However, to reiterate, we do not expect this to be the case based on recent theoretical developments; instead, we expect all messages to have the same unintended

consequence because they prime beliefs about the endemic nature of corruption. It is worth noting that if this expectation is born out, and there is no substantively different impact between the messages even though *one addresses taxation directly*, it would represent particularly strong evidence that the outcomes we see are due to general corruption priming effects.

3.3 | Dependent variable

To measure tax morale, our survey asked for respondents' agreement with the statement: "*citizens have a duty to pay tax to any government, no matter how the government performs*". Twenty-six percent of our full sample strongly agreed with this notion, while a further 50% indicated that they simply agreed. Only 4% strongly disagreed and 10% simply disagreed. The remaining 10% neither disagreed nor agreed.

Whenever survey data is used to investigate an issue such as commitment to tax payment there is a question as to whether self-reported beliefs reflect actual behavior. This is especially the case when it comes to tax payment, because failing to pay taxes is illegal and so citizens may feel the need to obscure their real-world actions. For this reason, we do not ask whether an individual has paid their taxes in the past year, but rather whether "*citizens have a duty to pay tax to any government*". This question allows respondents to answer honestly without implicating themselves, and so is likely to be more accurate. Significantly, Bodea and LeBas (2016, p. 222–223) argue that this form of question is "both consistent with other authors" use and "avoids some potential pitfalls"'. They also conclude that this wording provides a more reliable guide as to "societal attitudes toward taxation and actual evasion behavior" than other kinds of questions, such as the World Values Survey, which asks respondents how justifiable it is to cheat when it comes to tax payment, and as a result "over reports levels of tax morale in Nigeria".

4 | ESTIMATION STRATEGY

To examine the influence of messaging in a study like this, pair-wise difference in means (DIM) tests are appropriate to use *if* the only differences between treatment groupings are that they were exposed to different treatments or no treatment at all. To determine this, DIM tests were run on basic demographic variables. The results showed that there were likely important differences between some of the groups with respect to their mean levels of poverty, but not with respect to their mean levels of education, gender distribution or mean age.[2] On this basis, instead of using DIM tests, we ran a multivariate regression analysis, wherein assignments to treatment groups are treated as independent variables of interest (the baseline is the *control group*), and poverty is controlled for; this enables us to ensure that varying responses to the treatments across groups were not due to groups' different mean poverty levels. For details of how poverty was measured see Appendix C.

Responses to the duty to pay taxes survey question were recorded along a 5-point Likert agreement scale, but for ease of exposition we transform them into dichotomous categories (agreement or not). Doing this allows us to report a logistic regression model, rather than an ordered logistic regression model, the results of which have a more straightforward and intuitive interpretation. We also report the predicted probability shifts associated with exposure to the messages because they have a straightforward interpretation. They show the direction of association between exposure to messages and agreement with our dependent variable statement.

A positive shift illustrates that exposure to a message increases the likelihood of agreement that citizens always have a duty to pay tax—and they also indicate the substantive size of the effect that exposure generates.

5 | HOW DO (ANTI)CORRUPTION MESSAGES INFLUENCE WILLINGNESS TO PAY TAX?

We find considerable support for the expectation that exposure to awareness raising reduces agreement that citizens should pay taxes. The results in Table 1 show that exposure to four of the five messages is associated with a reduction in agreement that citizens should pay taxes, regardless of government performance. Specifically, the impacts of exposure to messaging are "significant" at a two-tailed p-value threshold of less than 0.05 for the *government success (p-value: 0.025), local (p-value: 0.049)*, and *tax* treatments (p-value: 0.040), and as this hypothesis is unidirectional, we also recognize the *widespread* message as being marginally "significant" (one-tailed, p-value: 0.055).

Interestingly, the estimated sizes of the effects that exposure to each of these four treatments has on agreement that citizens should pay taxes are strikingly similar. The estimated shifts in predicted probabilities show that, compared to the control group, exposure to both the *widespread* and *government success* treatments reduced the probability of agreement that citizens should pay taxes by seven percentage points, while exposure to the *local* and *tax* treatments reduced it by six percentage points. These are substantively large shifts; to give some context, exposure to messages about corruption and anticorruption were previously estimated to increase the probability of paying a bribe by an average of 6.5% points (Cheeseman and Peiffer 2022; Table 1). It is

T A B L E 1 How messaging influences agreement that citizens should pay taxes, regardless of government performance

	b	S.E.	Δ P.P.
Treatment groups			
Widespread	−0.26+	0.16	−0.07
Religious	−0.09	0.17	−0.01
Gov't success	−0.36**	0.16	−0.07
Local	−0.32**	0.16	−0.06
Tax	−0.33**	0.26	−0.06
Control			
Poverty	−0.23**	0.05	
Constant	1.62**	0.14	
N	2543		
Pseudo R^2	0.01		
Likelihood ratio	27.07		
Prob > Chi^2	0.000		

Note: Displayed are coefficients and standard errors (S.E.), with symbols indicating p-value thresholds: ** two-tailed p-value <.05, & + one-tailed p-value <.10. To estimate the substantive size of the influence of exposure to the messages, we report predicted probability (P.P.) shifts in the final column. These shifts were calculated from post-estimated analyses using marginal effects in Stata, where the effects of other variables in the model were held constant.

important to note too that the findings in Table 1 are robust to different specifications (found in Appendix C).[3]

Additional Wald tests were conducted to examine whether exposure to these four messages—*widespread, government success, local* and *taxes*—influenced agreement to the same degree. Specifically, these tests examine whether the coefficients reported in Table 1 associated with each of these four treatment groups are likely statistically different from each other. Table 2 shows the results of six Wald tests with each test's estimated chi-squared and probability of chi-squared, in parentheses; one test was run for each potential comparison between these four treatment groups. None of the probabilities of chi-squared are less than 0.10, indicating that there is likely no important statistical differences in the estimated impacts of exposure between the four treatments.[4] Put differently, these additional tests support the notion that all four messages negatively influence agreement that citizens should pay taxes to the same degree.

That these very different messages are impactful in the same way and to the same extent lends considerable support to the notion that different anticorruption messaging can have an equally negative influence on attitudes toward paying taxes. This finding is analogous to Peiffer's (2018) finding that four different messages had the *same* impact on concerns about corruption and beliefs that ordinary people could fight corruption. That very different messages about corruption can have the same influence makes sense in light of the lessons of the political psychology and cognitive response research reviewed above. For issues like corruption, people often discard or pay less attention to information in messages that does not confirm strongly held pre-existing attitudes (Lang, 2000; Lenz, 2009; Pfau et al., 1997). In our case, respondents appear to have paid less attention to the different tones and contents of these four messages, and instead were similarly primed by their core topic to recall pre-existing beliefs that corruption is systemic.

The exception in Table 1 is, of course, the impact of exposure to the *religious* treatment, which is not found to have a likely impact on agreement that citizens should pay tax. This is notable because it suggests that there may be at least one type of message about corruption that does not negatively impact attitudes about paying taxes—at least in the case of Lagos.

Why is this the case? Of all the treatments that we tested, the *religious* message is the only one that does not make any mention of public officials, political leaders or state institutions. There is an allusion to "those who use their positions and opportunities to cheat and loot", but nothing specific is said about who these people are. While more research is needed to test whether this finding holds in different kinds of contexts, one potential explanation is that when it comes to the effect of anti-corruption messages on tax morale, it is not solely priming citizens to think about corruption that generates negative externalities for attitudes toward taxation. Rather, precisely because taxes are paid to the state, what matters is priming individuals to think about corruption while simultaneously priming them to think about the political system—whether in terms of individual leaders, civil servants or state bodies. This makes intuitive sense. While "apolitical" messages may encourage respondents to focus on the problem of corruption, they may not

TABLE 2 Are the impacts of different messaging different?

	Results of coefficient comparison *t*-tests		
	Gov't success	**Local**	**Tax**
Widespread	0.40 (0.52)	0.13 (0.72)	0.20 (0.65)
Tax	0.04 (0.85)	0.01 (0.92)	
Local	0.08 (0.77)		

Note: Displayed are the chi² and the Prob > chi² (in parentheses) of six Wald tests.

encourage individuals to view it as a direct consequence of government failure. By contrast, the other "political" messages give citizens a clearer indication of who is to blame, and in priming both corruption and the state, undermine tax morale.

This interpretation receives strong support from the literature on the priming of public attitudes toward the performance and merits of state agencies. As Caillier notes (2018, p. 203–204), a growing number of studies demonstrate that when "negative reports reinforce undesirable stereotypes about [government] agencies, citizens will lower their evaluation of these organizations". In many cases, this effect appears to work through the same kind of subconscious priming posited here. Significantly, studies conducted in a range of other contexts have found that criticism of government bodies and commentary about corruption undermines confidence in the effectiveness of the state, with negative consequences for evaluations of politicians, government departments, and public agencies (Faulkner et al., 2015; Hvidman & Andersen, 2016).

This explanation appears to be more plausible than the main alternative, namely, that individuals are responding to something distinctive about the religious message such as the greater moral authority of religious leaders, which could possibly have had the distinctive effect of encouraging respondents to think about corruption and taxes in moral terms (i.e., that corruption is morally wrong). However, this is unlikely to be the case, because exposure to the religious treatment was not associated with greater agreement that corruption is immoral or that reporting corruption is the morally right thing to do in our survey.[5] Moreover, the same message has previously been found to have increased readers' willingness to pay a bribe (Cheeseman and Peiffer 2022), which is strong evidence that the authority of religious leaders is not sufficiently powerful to counteract the priming effects of anti-corruption messaging. Given this, our results are much more likely to be explained by the interaction of state and corruption priming effects—and the lack of this interaction in the case of the religious treatment—the implications of which we discuss at greater length in the conclusion.[6]

6 | CONCLUSION: THE UNINTENDED CONSEQUENCES OF AWARENESS RAISING CAMPAIGNS

This article has assessed whether anticorruption campaigns have unintended—and unwanted—consequences for the social contract. Testing the impact of five messages in Lagos, Nigeria, we found that none of our anti-corruption messages had a positive effect on attitudes toward taxation. Moreover, all but one treatment led to a fall in tax morale. It is important to avoid simply assuming that all individuals who become less willing to pay taxes subsequently pay less tax—some may be forced to pay via automatic deductions from their payslip, while others may continue to pay because they are afraid of the potential penalties for tax avoidance. On average, however, it is likely that a significant decline in tax morale will negatively impact on tax payment in a context in which such a large proportion of economic activity is informal and state capacity to enforce payment has been historically weak. This is certainly the belief of the Lagos State Government, which has invested considerable resources in public relations campaigns to boost tax morale precisely because it believes that this is critical to closing the taxation gap (Cheeseman & de Grammont, 2017).

Given the importance of taxes for government revenue and the relationship between citizens and the state, this finding is just as significant as the tendency for anti-corruption messaging to foster "corruption fatigue" and hence induce bribe payment (Cheeseman & Peiffer, 2020; Corbacho et al., 2016). This argument does not imply that citizens are doing anything wrong

if they reduce their commitment to tax payment after receiving an anti-corruption message—in many ways, it is a rational response to the information that they have been presented with. But given the fact that such messaging often leads citizens to overestimate the true extent of corruption (Cheeseman & Peiffer, 2020), and the importance of the social contract not just for tax payment but also citizen compliance in other areas, the consequences of this process are clearly significant for wider efforts to build more effective states.

How far can this finding, based on a study of Lagos, be generalized? There are a number of features of Lagosian politics that could be argued to render the association between politicians, government and corruption particularly salient and powerful. Although the effectiveness of the Lagos State Government is widely recognized to have improved under Governors Bola Tinubu (1999–2007) and Fashola (2007–2015), with Lagos becoming increasingly financially independent of the central government during their tenure (Bodea & LeBas, 2016), both have been accused of misusing state funds (Akinyole, 2015). Moreover, Tinubu is a notorious "godfather" (Majekodunmi & Awosika, 2013) who used his wealth to strengthen his political control, and his strategy for transforming Lagos was not to pursue "good governance" but rather to develop a productive form of clientelism and patronage politics in which key figures stood to gain —both politically and personally—from Lagos increasing its tax payment (Cheeseman & de Grammont, 2017). Lagosians have also witnessed a series of poorly performing governments at the national level, with notorious examples of grand corruption (Nwozor et al., 2020).

Yet while popular concern with corruption and nepotism has been shown to be particularly high in Nigeria, it is clear from the work of Caillier (2018) and Hvidman and Andersen (2016) that popular skepticism regarding the role and effectiveness of government is pronounced and impactful in very different contexts. Indeed, the fact that powerful priming effects have been detected in Denmark (Hvidman & Andersen, 2016), which Francis Fukuyama has held up as an example of a capable and law-bound state, suggests that a similar relationship between anti-corruption messaging and belief in tax payment may hold even in contexts that feature higher levels of democracy and lower corruption.

It is therefore even more noteworthy that we find one exception to our general rule: the religious treatment did not have a negative effect on tax payment. While more research needs to be conducted on this relationship, the main reason for the different outcome appears to be that the religious treatment was the only one that did not explicitly mention state officials, institutions and political leaders. This suggests a clear way to target anti-corruption messages that may minimize their impact on attitudes toward the social contract and the political system. If further research substantiates this finding, designing messages that do not prime individuals to think about the government/the state may be a viable option. For while the religious treatment did not name political leaders and institutions, it did describe "those who use their positions and opportunities to cheat and loot". It may therefore be possible to create a message that conveys the need to hold those in power accountable without priming citizens to always think about the state as part of the problem rather than part of the solution.

However, this "solution" is not without its own challenges. Designing anti-corruption drives that do not mention political institutions or leaders would clearly be challenging and many anti-corruption activists may feel that it would be disingenuous to design campaigns that intentionally avoid mentioning the worst offenders. Moreover, we also need to understand how to develop messages that do not just avoid negative effects but rather have strong positives ones. This may require communicating injunctive norms that describe the strength of public feeling against rule breaking in a given issue (Cheeseman and Peiffer 2022), emphasizing not current behavior but popular aspiration. Yet even cleverly designed messages will need to be carefully

tested. Where dealing with pervasive "social bads" are concerned, the challenge of "doing no harm" is even greater than it first appears.

CONFLICT OF INTEREST
Neither author has a conflict of interests. We do not work for, profit from, or have a stake in the organizations and processes discussed in the article.

DATA AVAILABILITY STATEMENT
Our data will be made publicly available prior to the publication of this article to support and promote data sharing. The research was completed relatively recently and the process of preparing and submitting the data is underway. We are committed to sharing the data and are very happy to use whatever platform you recommend, if Governance has a preferred option.

ENDNOTES
[1] A pilot study was conducted to ensure that the messages were well understood by the public. For a longer discussion of the gestation of the treatments see (Cheeseman and Peiffer 2022).

[2] The results showed that the *control* and *local* group's mean poverty score were likely significantly lower than that of the *religious* group (*t*-value for difference between *local* and *religious* groups was: 0.063, and *t*-value for difference between *control* and *religious* groups was 0.079; *t*-values for all other differences were over 0.10).

[3] This includes DIMs tests and a model in which we control for demographic variables, as well as whether the respondent has witnessed corruption or is a ruling party supporter. These results are consistent with the results reported in Table 1. Additionally, we also test whether the findings are robust in ordered logistic regression models where the dependent variable is measured using 3- and 5- point scales. In the case of the former, the results are robust to the findings in Table 1. In the case of the latter, the results are robust for all treatments bar the tax treatment (model 2, Table C.1, Appendix C). However, less than 10% of our respondents indicated a neutral position to our dependent variable question and half of the sample indicated simply "agree." According to Marks et al. (2012), this type of distribution justifies dichotomizing a variable's response patterns because if treated as continuous, analyses may hide important associations. Importantly, for our purposes, these analyses confirm that, while exposure to the tax treatment *may not have encouraged strong disagreement*, the tax treatment does decrease the likelihood of agreement with the idea that citizens should pay taxes (Table 1 and model 3, Table C.1, Appendix C).

[4] A single Wald test was also run to test whether there are meaningful differences among all four coefficients, the results of which also support the findings of the six tests reported in Table 2 (chi²: 0.43, Prob > chi²: 0.93).

[5] We ran two ordered logistic models, with dependent variables, respectively measuring agreement on a 5-point scale with the following statements: "if something corrupt is done for the right reasons, it isn't morally wrong" and "I would report corruption because it's the morally right thing to do". Like in Table 1, we examined whether exposure messaging was associated with agreement with both statements and controlled for the potential impact of poverty. In both tests, the *p*-value associated with exposure to the religious treatment was over 0.300, indicating that it does not shape (dis)agreement.

[6] A further alternative explanation is that our messages are simply capturing a "state priming" effect rather than an "anti-corruption" one. However, the fact that the religious treatment has previously been found to increase bribe payment, even though the state itself is not mentioned, demonstrates that corruption messages have a substantive effect independent of any "state" effect.

REFERENCES

Abbink, J. (2009). The Ethiopian second republic and the fragile "social contract". *Africa Spectrum*, *44*(2), 3–28. https://doi.org/10.1177/000203970904400201

Afrobarometer. (2018). *Summary of results, Afrobarometer round 7, survey in Nigeria, 2017*. Afrobarometer. Available from http://afrobarometer.org/sites/default/files/publications/Summary%20of%20results/nig_r7_sor_28032018_eng.pdf

Afrobarometer. (2020). Surveys and methods. Available from https://www.afrobarometer.org/surveys-and-methods

Akinloye, L. (2015, September 8). Nigeria's political golden boy facing allegations of corruption. *The Guardian*. Available online https://www.theguardian.com/world/2015/sep/08/nigeria-corruption-babatunde-fashola-allegations

Andersson, S., & Heywood, P. M. (2012). Anti-corruption as a risk to democracy: on the unintended consequences of international anti-corruption campaigns. In L. De Sousa, B. Hindess & P. Larmour (Eds.), Governments, NGOs and Anti-Corruption. Oxford: Routledge, pp. 59–76.

Berinsky, A., Hutchings, V., Mendelberg, T., Lee, S., & Valentino, N. (2011). Sex and race: Are black candidates more likely to be disadvantaged by sex scandals? *Political Behavior*, *33*(2), 179–202. https://doi.org/10.1007/s11109-010-9135-8

Bodea, C., & LeBas, A. (2016). The origins of voluntary compliance: Attitudes toward taxation in urban Nigeria. *British Journal of Political Science*, *46*(1), 215–238. https://doi.org/10.1017/s000712341400026x

Brody, R., & Page, B. (1972). Comment: Assessment of policy voting. *American Political Science Review*, *66*(2), 450–458. https://doi.org/10.2307/1957788

Burbidge, D. (2016). *The shadow of Kenyan democracy: Widespread expectations of widespread corruption*. Routledge.

Cacioppo, J. T., & Petty, R. E. (1989). Effects of message repetition on argument processing, recall, and persuasion. *Basic and Applied Social Psychology*, *10*(1), 3–12. https://doi.org/10.1207/s15324834basp1001_2

Caillier, J. (2018). The priming effect of corruption and bureaucracy bashing on citizens' perceptions of an agency's performance. *Public Performance & Management Review*, *41*(2), 201–223. https://doi.org/10.1080/15309576.2018.1431138

Carmines, E., & Stimson, J. (1989). *Issue evolution: Race and the transformation of American politics*. Princeton University Press.

Carter, T., Ferguson, M., & Ran, H. (2011). A single exposure to the American flag shifts support toward republicans up to 8 months later. *Psychological Science*, *22*(8), 1011–1018. https://doi.org/10.1177/0956797611414726

Cheeseman, N., & de Gramont, D. (2017). Managing a mega-city: Learning the lessons from Lagos. *Oxford Review of Economic Policy*, *33*(3), 457–477. https://doi.org/10.1093/oxrep/grx033

Cheeseman, N., & Peiffer, C. (2020). *The unintended consequences of anti-corruption messaging in Nigeria: Why pessimists are always disappointed*. SOAS ACE working paper 024. SOAS University of London. Available from https://ace.soas.ac.uk/wp-content/uploads/2020/06/ACE-WorkingPaper024-NigeriaUnintendedConsequences-200608.pdf

Cheeseman, N., & Peiffer, C. (2021). The curse of good intentions: Why anticorruption messaging can encourage bribery. *American Political Science Review*, *116*(3), 1–15. https://doi.org/10.1017/S0003055421001398

Cheeseman, N., & Peiffer, C. (2022). The Curse of Good Intentions: Why Anticorruption Messaging Can Encourage Bribery. *American Political Science Review*, *116*(3), 1081–1095.

Corbacho, A., Gingerich, D., Oliveros, V., & Ruiz-Vega, M. (2016). Corruption as a self-fulfilling prophecy: Evidence from a survey experiment in Costa Rica. *American Journal of Political Science*, *60*(4), 1077–1092. https://doi.org/10.1111/ajps.12244

D'Arcy, M. (2011). *Why do citizens assent to pay tax? Legitimacy, taxation and the African state?* Afrobarometer Working Paper No. 126. Afrobarometer. Available from http://afrobarometer.org/publications/wp126-why-do-citizens-assent-pay-tax-legitimacy-taxation-and-african-state

Devarajan, S., & Ianchovichina, E. (2017). A broken social contract, not high inequality, led to the Arab Spring. *Review of Income and Wealth*, *64*(1), 5–25. https://doi.org/10.1111/roiw.12288

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt Brace Jovanovich College Publishers.

Faulkner, N., Martin, A., & Peyton, K. (2015). Priming political trust: Evidence from an experiment. *Australian Journal of Political Science*, *50*(1), 164–173. https://doi.org/10.1080/10361146.2014.979759

Gadenne, L. (2017). Tax me, but spend wisely? Sources of public finance and government accountability. *American Economic Journal: Applied Economics*, *9*(1), 274–314. https://doi.org/10.1257/app.20150509

Hampton, J. (1980). Contracts and choices: Does Rawls have a social contract theory? *The Journal of Philosophy*, *77*(6), 315–338. https://doi.org/10.2307/2025640

Hayek, F. A. (2014). *The road to serfdom: Text and documents: The definitive edition*. Routledge.

Hickey, S. (2011). The politics of social protection: What do we get from a 'social contract' approach? *Canadian Journal of Development Studies*, *32*(4), 426–438. https://doi.org/10.1080/02255189.2011.647447

Hobbes, T. (1651). *Leviathan: Or the matter, forme and power of a commonwealth, ecclesiasticall and civil*. Andrew Crooke.

Hoffman, L. K., & Patel, R. N. (2017). *Collective action on corruption in Nigeria: A social norms approach to connecting society and institutions*. Chatham House Report, May. Chatham House. Available from https://reader.chathamhouse.org/collective%2Daction%2Dcorruption%2Dnigeria%2Dsocial%2Dnorms%23

Hvidman, U., & Andersen, S. C. (2016). Perceptions of public and private performance: Evidence from a survey experiment. *Public Administration Review*, *76*(1), 111–120. https://doi.org/10.1111/puar.12441

Igiebor, G. O. (2019). Political corruption in Nigeria: Implications for economic development in the Fourth Republic. *Journal of Developing Societies*, *35*(4), 493–513. https://doi.org/10.1177/0169796x19890745

Jahnke, B., & Weisser, R. A. (2019). How does petty corruption affect tax morale in sub-Saharan Africa? *European Journal of Political Economy*, *60*, 101751. https://doi.org/10.1016/j.ejpoleco.2018.09.003

Kobis, N., Troost, M., Brandt, C., & Soraperra, I. (2019). Social norms of corruption in the field: Social nudges on posters can help to reduce bribery. *Behavioural Public Policy*, 1–28. https://doi.org/10.1017/bpp.2019.37

Lang, A. (2000). The limited capacity model of mediated message processing. *Journal of Communication*, *50*(1), 46–70. https://doi.org/10.1111/j.1460-2466.2000.tb02833.x

Lawal, G. (2007). Corruption and development in Africa: Challenges for political and economic change. *Humanity & Social Sciences Journal*, *2*(1), 1–7.

Lenz, G. (2009). Learning and opinion change, not priming: Reconsidering the priming hypothesis. *American Journal of Political Science*, *53*(4), 821–837. https://doi.org/10.1111/j.1540-5907.2009.00403.x

Majekodunmi, A., & Awosika, F. O. (2013). Godfatherism and political conflicts in Nigeria: The Fourth Republic in perspective. *International Journal of Management and Social Sciences Research*, *2*(7), 70–75.

Mangoting, Y., Sukoharsono, E. G. (2015). Developing a model of tax compliance from social contract perspective: Mitigating the tax evasion. *Procedia – Social and Behavioral Sciences*, *211*, 966–971. https://doi.org/10.1016/j.sbspro.2015.11.128

Marks, G., Steenbergen, M., & Hooghe, L. (2012). To dichotomize or not to dichotomize? *European Union Politics*, *13*(2), 334–339. https://doi.org/10.1177/1465116511435823

Marquette, H. (2012). "Finding God" or "Moral disengagement" in the fight against corruption in developing countries? Evidence from India and Nigeria. *Public Administration and Development*, *32*(1), 11–26. https://doi.org/10.1002/pad.1605

McCulloch, N., Moerenhout, T., & Yang, J. (2021). Building a social contract? Understanding tax morale in Nigeria. *The Journal of Development Studies*, *57*(2), 226–243. https://doi.org/10.1080/00220388.2020.1797688

Meagher, K. (2018). Taxing times: Taxation, divided societies and the informal economy in Northern Nigeria. *The Journal of Development Studies*, *54*(1), 1–17. https://doi.org/10.1080/00220388.2016.1262026

Moore, M., Prichard, W., & Fjeldstad, O. H. (2018). *Taxing Africa: Coercion, reform and development*. Zed Books Ltd.

Ndinga-Kanga, M., van der Merwe, H., & Hartford, D. (2020). Forging a resilient social contract in South Africa: Societies sustaining peace in the post-apartheid era. *Journal of Intervention and Statebuilding*, *14*(1), 22–41. https://doi.org/10.1080/17502977.2019.1706436

Nwozor, A., Olanrewaju, J. S., Oshewolo, S., & Ake, M. B. (2020). Is Nigeria really fighting to win the anti-corruption war?: Presidential body language, "string-puppetting" and selective prosecutions. *Journal of Financial Crime*, *27*(2), 601–617. https://doi.org/10.1108/jfc-08-2019-0109

Ogundiya, I. S. (2017). Political corruption in Nigeria: Theoretical perspectives and some explanations. *The Anthropologist*, *11*(4), 281–292. https://doi.org/10.1080/09720073.2009.11891117

Ortega, D., Ronconi, L., & Sanguinetti, P. (2016). Reciprocity and willingness to pay taxes: evidence from a survey experiment in Latin America, *Economía*, *16*(2), 55–87.

Paler, L. (2013). Keeping the public purse: An experiment in windfalls, taxes, and the incentives to restrain government. *American Political Science Review*, *107*(4), 706–725. https://doi.org/10.1017/s0003055413000415

Peiffer, C. (2017). *Getting the message: Examining the intended – And unintended – Impacts of corruption awareness-raising*. Developmental Leadership Program. Available from https://www.dlprog.org/publications/research-papers/getting-the-message-examining-the-intended-and-unintended-impacts-of-corruption-awareness-raising

Peiffer, C. (2018). Message received? Experimental findings on how messages about corruption shape perceptions. *British Journal of Political Science*, *50*(3), 1207–1215. https://doi.org/10.1017/s0007123418000108

Peiffer, C., & Walton, G. (2019). *Overcoming collective action problems through anti-corruption messages*. Development Policy Centre Discussion Paper No. 77. Development Policy Centre, Crawford School of Public Policy. Available from https://ssrn.com/abstract=3333475

Pfau, M., James Tusing, K., Koerner, A. F., Lee, W., Godbold, L. C., Penaloza, L. J., Shu-Huei Yang, V., & Hong, Y.-H. (1997). Enriching the inoculation construct: The role of critical components in the process of resistance. *Human Communication Research*, *24*(2), 187–215. https://doi.org/10.1111/j.1468-2958.1997.tb00413.x

Prichard, W. (2015). *Taxation, responsiveness and accountability in sub-Saharan Africa: The dynamics of tax bargaining*. Cambridge University Press.

Prichard, W. (2019). Tax, politics and the social contract in Africa. In N. Cheeseman (Ed.), *Oxford research encyclopaedia, politics*. Oxford University Press. https://doi.org/10.1093/acrefore/9780190228637.013.853

Rawls, J. (1971). *A theory of justice*. Harvard University Press.

Riker, W. (1986). *The art of political manipulation*. Yale University Press.

Rousseau, J. J. (2018). *Rousseau: The social contract and other later political writings*. Cambridge University Press.

Smith, D. J. (2008). *A culture of corruption: Everyday deception and popular discontent in Nigeria*. Princeton University Press.

Tilly, C. (1990). *Coercion, capital, and European states, AD 990–1992*. Basil Blackwell.

Transparency International. (2019). *Corruption perception index data for Nigeria*. Transparency International. Available June 6, 2020 from https://www.transparency.org/en/countries/nigeria

Transparency International. (2020). *Corruption perceptions index*. Transparency International. Available from https://www.transparency.org/en/cpi/2019

Umar, M. A., Derashid, C., Ibrahim, I., & Bidin, Z. (2018). Public governance quality and tax compliance behavior in developing countries: The mediating role of socioeconomic conditions. *International Journal of Social Economics*, *46*(3), 338–351. https://doi.org/10.1108/ijse-11-2016-0338

United Nations. (2004). *United Nations convention against corruption*. United Nations Office on Drugs and Crime. Available from https://www.unodc.org/documents/treaties/UNCAC/Publications/Convention/08-50026_E.pdf

Uslaner, E. M. (2010). Tax evasion, corruption, and the social contract in transition. In J. Alm, J. Martinez-Vazquez, & B. Torgler (Eds.), *Developing alternative frameworks for explaining tax compliance* (pp. 174–190). Routledge.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Cheeseman, N., & Peiffer, C. (2022). Why efforts to fight corruption can undermine the social contract: Lessons from a survey experiment in Nigeria. *Governance*, 1–17. https://doi.org/10.1111/gove.12720