AJPS AMERICAN JOURNAL of POLITICAL SCIENCE

Corruption as a Self-Fulfilling Prophecy: Evidence from a Survey Experiment in Costa Rica

Ana Corbacho International Monetary Fund
Daniel W. Gingerich University of Virginia
Virginia Oliveros Tulane University
Mauricio Ruiz-Vega Macroeconomic Section-US Embassy in Jakarta, Indonesia

Abstract: *An influential literature argues that corruption behaves as a self-fulfilling prophecy. Its central claim is that the individual returns to corruption are a function of the perceived corruptibility of the other members of society. Empirically, this implies that if one were to exogenously increase beliefs about societal levels of corruption, willingness to engage in corruption should also increase. We evaluate this implication by utilizing an information experiment embedded in a large-scale household survey recently conducted in the Gran Área Metropolitana of Costa Rica. Changes in beliefs about corruption were induced via the random assignment of an informational display depicting the increasing percentage of Costa Ricans who have personally witnessed an act of corruption. Consistent with the self-fulfilling prophecy hypothesis, we find that internalizing the information from the display on average increased the probability that a respondent would be willing to bribe a police officer by approximately .05 to .10.*

Replication Materials: The data, code, and any additional materials required to replicate all analyses in this article are available on the *American Journal of Political Science* Dataverse within the Harvard Dataverse Network, at: http://dx.doi.org/10.7910/DVN/8GEYKS.

A question of enduring interest in the social sciences is under what conditions individuals decide to engage in illicit and socially harmful forms of behavior. Responses to this question vary widely, but one can organize thinking on this issue into two groups of arguments based on the role they assign to an individual's community environment. The role of community may be analytically relegated to the background, viewed as investing individuals with norms and a sanction schedule that "prices" illicit behavior in particular ways, but otherwise not operating directly on individual choices. Alternatively, the role of community may be placed at the forefront, such that discussions about the individual returns to illicit behavior are deemed to be largely meaningless without reference to the choices of other community members about engaging in similar behavior. The difference here is between viewing illicit behavior as an individual phenomenon embedded within a particular context and viewing illicit behavior as an intrinsically social phenomenon.

Both views are encountered in the contemporary literature on corruption, which is our focus here. A relatively new and growing empirical literature frames its analysis of corruption largely along the lines of the first view. An

Ana Corbacho is Division Chief, International Monetary Fund, 700 19th Street NW, Washington, DC 20431 (acorbacho@imf.org). Daniel W. Gingerich is Associate Professor and Director, Quantitative Collaborative, Woodrow Wilson Department of Politics, University of Virginia, P.O. Box 400787 Charlottesville, VA 22904-4787 (dwg4c@virginia.edu). Virginia Oliveros is Assistant Professor, Department of Political Science, Tulane University, 6823 St. Charles Avenue Norman Mayer Building, Room 311 New Orleans, LA 70118 (voliveros@tulane.edu). Mauricio Ruiz-Vega is a Development Consultant, Macroeconomic Section-US Embassy in Jakarta, Indonesia, Jalan Medan Merdeka Selataan, Jakarta (mau.ruiz54@gmail.com).

Thanks to the Editor, Sarah Hummel, Jason Lyall, Jack Paine, Peter van der Windt, Rebecca Weitz-Shapiro, and three anonymous reviewers for helpful comments. Participants at the 2015 American Political Science Association annual meeting, the 2015 European Political Science Association annual meeting, the 2015 Midwest Political Science Association annual meeting, and seminars at Duke University, Harvard University, the Inter-American Development Bank, and Vanderbilt University also provided valuable feedback. This article was prepared when Ana Corbacho was Sector Economic Advisor and Daniel Gingerich and Virginia Oliveros were visiting scholars at the Inter-American Development Bank. Research was conducted with generous support from the Inter-American Development Bank.

American Journal of Political Science, Vol. 60, No. 4, October 2016, Pp. 1077–1092

 DOI: 10.1111/ajps.12244

older but still vibrant game-theoretic literature explicitly adopts the second view, arguing that levels of corruption emerge endogenously from a society-wide coordination game in which the individual returns to corrupt behavior are increasing in the inclination toward corruption of the other members of society. The two approaches have distinctive implications. The first implies that holding individual moral values and expectations about punishment constant, beliefs about the societal frequency of corruption should hold little sway over decisions about corrupt action. The second view holds the opposite: The higher individuals perceive the level of corruption in society to be, the more inclined they will be to engage in corrupt behavior themselves. In this latter scenario, corruption behaves as a self-fulfilling prophecy.

Does corruption corrupt? Does the level of perceived corruption in a society affect an individual's willingness to engage in corrupt behavior? We argue in this article that the answer to this question is yes, that the decision to engage in corrupt behavior is crucially shaped by perceptions of what other actors in society are doing. In this sense, we argue that individuals approach the choice to engage in corruption much as a game theorist would, taking strategic interdependence into account in assessing the costs and benefits of their actions. In particular, using data from an original survey we conducted in Costa Rica, we show that learning about increasing levels of corruption in society increases the likelihood that citizens would be willing to bribe a police officer to avoid paying a traffic ticket.

Our conclusions are based on a carefully crafted empirical research design. To deal with issues of social desirability bias, we employ a novel technique that combines the bias-reducing advantages of sensitive survey techniques with direct questioning, which results in more precise estimates. To deal with the problem of confounding, we induce exogenous variation in beliefs about corruption using a survey experiment that provides information about the rising levels of corruption in Costa Rica. In this way, our article advances the literature on corruption in two ways. First, we provide the first experimental evidence about the effects of perceived corruption in society on an individual's willingness to engage in petty corruption. Although a number of formal papers have previously argued that corruption corrupts, our study is unique in providing convincing empirical evidence that this is actually the case. Second, we use a novel method to measure corruption that significantly reduces bias and increases precision—a technique that can easily be replicated in the study of other sensitive activities.

Costa Rica provides a propitious environment for studying corruption. Although corruption is relatively low by regional standards, the country recently experienced a substantial increase in perceived corruption. According to the 2013 Corruption Perception Index by Transparency International, Costa Rica is ranked 49/177, considerably above Honduras (140), Nicaragua (127), Guatemala (123), and Panama (102). In fact, the only two Latin American countries that perform better than Costa Rica on this metric are Uruguay (19) and Chile (22). However, according to a nationally representative survey conducted by *Latinobarómetro*, the number of people who have witnessed an act of corruption has increased from 16% in 2006 to 24% in 2011.[1] At the elite level, the country has seen repeated political corruption scandals over the last decade, including the indictment of three different past presidents for bribery (and the conviction of two of them), as well as a number of forced resignations among cabinet officials. Corruption scandals of such magnitude have no precedent in Costa Rica. Given these abrupt changes occurring in a country once characterized as the "Switzerland of Central America," beliefs about how deeply embedded corruption is in Costa Rican society are likely to be in flux for many individuals. For this reason, the country is a natural setting for exploring how information about the scope of corruption may drive corrupt behavior.

The Choice for Corruption: Two Views of Decision Making

Broadly speaking, there are two basic views of the decision process by which an individual, faced with the opportunity, chooses to engage in a corrupt act or refrains from doing so. The first view is a decision-theoretic one. [2] In this view, the choice to engage in corrupt behavior results from a fundamentally atomistic and societally noncontingent risk–return calculus. Presented with the opportunity to pursue an illicit action, an actor engages in an introspection exercise in which she considers the potential rewards and opportunity costs of corrupt action, her personal moral views on the subject, and the fixed, institutionally determined likelihood of detection, as well as the magnitude of the sanction such detection brings. A society-wide level of corruption percolates up directly from the results of many such introspection exercises. This decision-theoretic perspective is implied in a large

[1] This perception of increasing corruption was corroborated by the focus groups we conducted prior to the survey. See the supporting information for details.

[2] Here, we use the term *decision-theoretic* (as opposed to *game-theoretic*) to denote the analysis of decision making by individual agents whose decisions do not affect the returns to the decisions adopted by other agents.

body of empirical work that puts pride of place on the demographic characteristics and values of individuals in explanations of corrupt behavior. Recent studies emphasizing the explanatory role of gender (Dollar, Fisman, and Gatti 2001; Esarey and Chirillo 2013; Sung 2003; Swamy et al. 2001), age (Torgler and Valev 2010), cultural attitudes (Cameron et al. 2009; Fisman and Miguel 2007), and partisan preferences (Anderson and Tverdova 2003; Anduiza, Gallego, and Muñoz 2013) all fit within this rubric.

The second vision is a game-theoretic one. In this perspective, the choice of an actor to engage in corrupt behavior results from a fundamentally interdependent and societally contingent risk–return calculus. Presented with the same opportunity, an actor's introspection exercise incorporates all of the attributes and considerations described above, but now also hinges crucially on beliefs about what other actors, faced with similar decisions of their own, are likely to do. As before, a society-wide level of corruption percolates up from these introspection exercises; however, here the resolution of each such exercise is contingent on the resolution of the others. This second view of decision making typically holds that choice over corrupt acts is characterized by strategic complementarities. For our purposes, this means that the decisions actors make about engaging in corruption are complementary to one another in the sense that the expected return that any given actor associates with engaging in corruption is increasing with the expected number of other actors who do so. Consequently, all actors have a strong incentive to coordinate their behavior, be that partaking in corrupt action or abstaining from it.

The vast majority of theoretical work on corruption in the political economy tradition adopts a game-theoretic perspective with strategic complementarities.[3] Work in this vein has long emphasized that corruption is subject to coordination dilemmas and herd behavior, and, as a consequence, that high (or low) levels of corruption tend to feed upon themselves and persist over time. The specific mechanisms adduced to explain why this is so are many and varied.

Some accounts concentrate on how the existence of corruption undermines sanctioning mechanisms, thereby furthering the incidence of corruption in the first place (Andvig and Moene 1990; Cadot 1987; Lui 1986; Mishra 2006). Search costs have also been invoked to explain how corruption corrupts: The more certain are parties to a corrupt exchange about the corruptibility of their partners, the less costly it is to consummate a corrupt bargain (Ryvkin and Serra 2012; see also Andvig and Moene 1990). Recent work emphasizes the role of guilt aversion in generating corruption spillovers (Balafoutas 2011). Theoretical frameworks taking a long-term view have emphasized mechanisms such as the allocation of talent between productive activities and rent seeking (Açemoglu 1995), reputational lock-in for collectivities (Tirole 1996), the intergenerational transmission of cultural values (Hauk and Saez-Marti 2002), and imitative processes of strategy selection (Accinelli and Sánchez Carrera 2012). Finally, several explicitly political accounts of corruption have argued that corruption spillovers emerge via the selection mechanism determining who holds public office (Caselli and Morelli 2004; Dal Bó, Dal Bó, and Di Tella 2006) or by affecting the bargaining power of political machines vis-à-vis rank-and-file members in the bureaucracy (Gingerich 2009). In all of this work, society-wide corruption is envisioned as an inherently emergent phenomenon.

Both the decision-theoretic and game-theoretic approaches toward choice over corrupt acts are plausible on their face, as becomes clear when they are applied to police stops for traffic infractions, one of the most common arenas of petty corruption and the focus of this article. In such situations, the potential for a mutually beneficial transaction that defrauds the public fisc is clearly present. For example, the driver could pay the police officer a bribe in some amount lower than the official sanction for the infraction, and, in so doing, both agents would be better off than if they had followed the letter of the law. Alternatively, the officer could insinuate his willingness to accept a bribe in lieu of applying the sanction, again providing both actors with a financial benefit.

According to the decision-theoretic approach, a corrupt transaction would be consummated if both the driver and the police officer independently assessed that the financial returns to the exchange were sufficiently high, that their normative qualms—if any—were sufficiently minor, and that the risk of detection and sanction was sufficiently small. Seen in this light, the proportion of police stops in a polity that would result in corrupt transactions would be completely determined by the distribution of utilities generated by the pecuniary returns to the transaction, the distribution of moral tastes for or against corruption among drivers and officers, and the quality of institutions that monitor and sanction corruption.

A game-theoretic perspective would analyze the same situation differently. It might begin emphasizing that actually executing a corrupt transaction is not easy. For instance, if the driver offers a bribe to an unwilling police officer, she runs the risk of sanction for attempted bribery in addition to that for the initial infraction. Likewise, if the police officer begins the process of extorting a bribe

[3] See Aidt (2003) for a review of this literature.

from a driver strongly disposed against corruption, he runs the risk of being reported. Given the limited information that the driver and police officer can glean about each other during the traffic stop, each will have to make a decision to initiate or not initiate a corrupt transaction based largely upon their beliefs about what the typical driver and typical officer are likely to do in such a scenario. As a consequence, expectations of social behavior are now central to the choice problem of each agent. Any given driver will be more inclined to initiate a corrupt transaction the greater the proportion of corruptly inclined police officers she believes there to be, since—in the absence of detailed information about the particular officer with whom she is dealing—she views said officer as a random draw from the population of officers. Similarly, and for the same reason, any given police officer would be more inclined to extort a bribe the greater the proportion of corruptly inclined drivers he believes there to be. The coordination of beliefs about the likely actions of others is central in this perspective, and, as such, it plays a crucial role in determining the proportion of police stops that ultimately result in corrupt transactions. Pecuniary returns, tastes for corruption, and institutions all remain relevant, but they alone are far from decisive in determining the prevalence of corrupt transactions.

In order to properly consider the implications of a game-theoretic decision process marked by strategic complementarity for corruption during traffic stops, we developed a formal model specifically tailored to capture what we view as the central elements of such encounters: anonymity, two-sided uncertainty, and belief conditionality (for both agents) of the returns to initiating a corrupt exchange. All details of the game—players, actions and timing, informational conditions, and formal proofs—are provided in the supporting information.

The central intuitions of our game are apparent in Figure 1. The proportion of drivers willing to denote a disposition toward corruption during a traffic stop, labeled q, is represented by the solid line displayed on the x-axis. This proportion is a function of the (expected) share of police officers disposed toward corruption. The proportion of police officers willing to denote a disposition toward corruption, labeled p, is represented by the dashed line displayed on the y-axis. This proportion is a function of the (expected) share of drivers disposed toward corruption. The equilibria for the corruption game are the points of intersection between the two lines (the circles).

The first item to note is the fact that there are multiple equilibria. Indeed, there are three equilibria: a high corruption equilibrium (p_H^*, q_H^*) in which all drivers and police officers indicate a disposition toward corruption (all police stops result in a corrupt transaction), a low corruption equilibrium (p_L^*, q_L^*) in which no drivers and officers indicate a disposition toward corruption (no police stops result in a corrupt transaction), and an intermediate equilibrium (p_M^*, q_M^*) in which the proportions of drivers and officers who indicate a disposition to engage in corruption fall within an interval between zero and one (some police stops result in a corrupt transaction).

Yet not all the equilibria merit the same consideration. Of the three, only the low equilibrium and the high equilibrium are stable.[4] In particular, if drivers and police were to have beliefs about each other that were close to but slightly different from either of the extreme equilibria, (i.e., located somewhere in a small neighborhood around p_H^*, q_H^* or p_L^*, q_L^*), beliefs and behavior would dynamically adjust until the high or low equilibrium, respectively, was eventually reached. The same would not occur for the intermediate equilibrium. Thus, the two extreme equilibria are robust to small perturbations in beliefs, whereas the intermediate equilibrium is not.

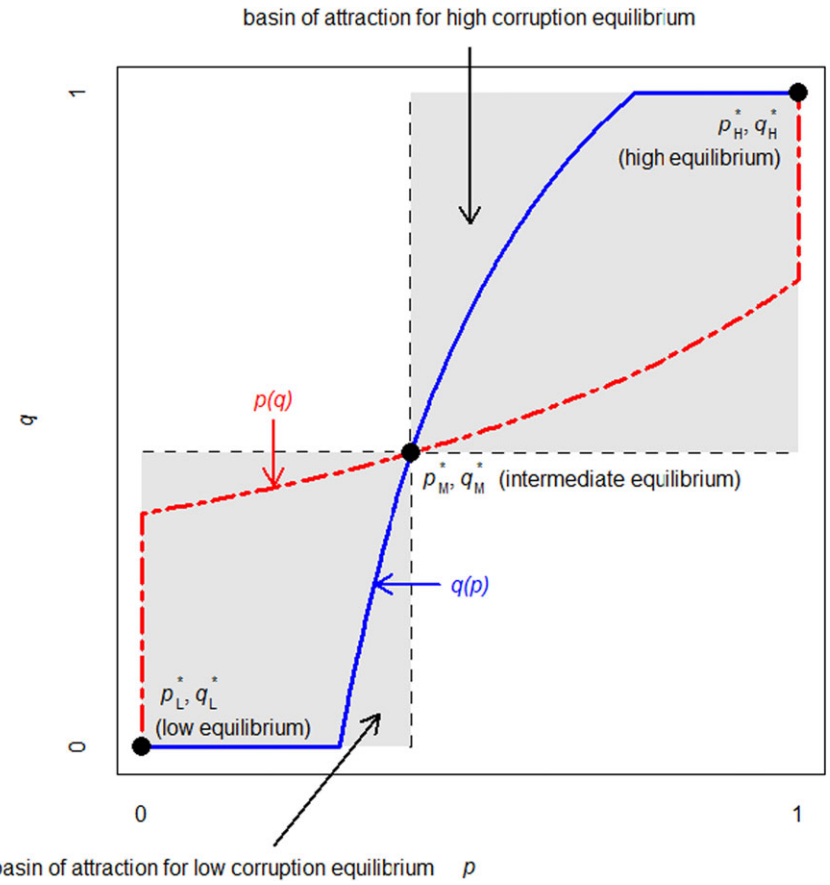
In this framework, pecuniary returns, moral tastes for corruption, and institutions shape the likelihood that one or the other of the two stable corruption equilibria will occur. More specifically, these items affect the relative size of the basins of attraction of the two equilibria (shaded in gray in Figure 1), defined as the set of initial beliefs that would ultimately lead a particular equilibrium to prevail. The severity of sanctions for corruption affects the relative size of the basins of the two plausible equilibria in a highly intuitive way: the greater the severity of sanctions, the larger (smaller) the relative size of the basin for the low (high) corruption equilibrium. Similarly intuitive is the influence of economic arrangements and cultural norms that determine tastes for corruption: the more intense said tastes, the smaller (larger) the relative size of the basin for the low (high) corruption equilibrium. Thus, the quality of institutions and moral tastes shape the scope of the gravitational pull of each plausible equilibrium, in so doing making driver–police coordination around one equilibrium point more or less likely than coordination around the other.[5]

However, the relevance of moral tastes and institutions notwithstanding, the model implies that the influence of these factors is subsidiary to the role of expectations. Even in a polity whose institutions strongly sanction corruption and whose citizenry finds it distasteful,

[4]To be specific, these equilibria satisfy a criterion for plausibility called dynamical stability; the intermediate equilibrium does not satisfy this criterion. See the supporting information for details.

[5]See Medina (2007) for a discussion of the probability of equilibria in coordination games.

FIGURE 1 Equilibria for the Corruption Game



Note: The proportion of citizens with a disposition toward corruption is represented by the solid line. The proportion of officials with a disposition toward corruption is represented by the dashed line. The equilibria for the corruption game are the points of intersection between the two lines (the circles).

sufficiently cynical beliefs could lead drivers and police to coordinate around the high corruption outcome. Similarly, a polity whose institutions are highly permissive to corruption and whose citizens are generally tolerant of it may still wind up on the low corruption equilibrium should drivers' and police officers' beliefs about one another be sufficiently sanguine. Expectations about corruption act as self-fulfilling prophecies. As such, constitutive beliefs about the typical behavior of members of society play a central explanatory role in game-theoretic models of corruption with strategic complementarities; they play no such role in analogous decision-theoretic frameworks.

It would be hard to overstate how different the policy implications of these two views are. From a decision-theoretic perspective, polities that have high levels of corruption are the way they are because economic arrangements generate high pecuniary returns to corrupt transactions, citizens and officials have "bad" preferences, institutions that monitor and sanction corruption are ineffective, or some combination of the above. Improving any one of these items will directly reduce the incidence of corruption. From a game-theoretic perspective, polities suffering from high levels of corruption may very well have "good" preferences and institutions but are the way they are because citizens and officials have coordinated around a set of highly pessimistic beliefs about one another. Changing the underlying fundamentals without altering the coordination of beliefs may not solve the problem.

In this article, we provide an explicit empirical evaluation of the relevance of beliefs about the incidence of corruption in society for individual choices about engaging in corrupt behavior. Our analysis clearly demonstrates that the inclination to act corruptly is contingent on beliefs about the scope of corruption in society. For nearly 30 years, game-theoretic models of corruption with strategic complementarities have emphasized the

importance of this belief contingency at a theoretical level. Our article is the first to empirically establish its relevance using experimental evidence.

Empirical Studies of Corruption Spillovers

The empirical literature examining the degree to which corrupt behavior exhibits strategic complementarities is still at a fairly early stage. Several papers have used aggregate data to show that corruption might be contagious (an empirical implication of complementarity). For example, Dong and Torgler (2012) use province-level data for China from 1998 to 2007 to show that social interaction has a significant positive effect on corruption. Using state-level U.S. data from 1995 to 2004, Goel and Nelson (2007) also find evidence of corruption neighborhood effects. Lopez-Valcarcel, Jiménez, and Perdiguero (2014) similarly find evidence that corruption is contagious using a data set of local Spanish municipalities from 2001 to 2010. Studies by Becker, Egger, and Seidel (2009) and Goel and Saunoris (2014) utilize cross-national data to estimate the degree to which corruption in one country affects its neighbors. In both cases, the authors find evidence of spillover effects. Contrary to the findings of these studies, Márquez, Salinas-Jiménez, and Salinas-Jiménez (2011) find no evidence of corruption spillovers.

Articles that explore the contagiousness or self-fulfilling prophecy hypothesis using individual-level data as we do in this article are few. Using data from the European Values Survey, Dong, Dulleck, and Torgler (2012) show that the more respondents perceive others as being corrupt, the more tolerant they are toward bribery. Similarly, a recent report from the Latin American Public Opinion Project (LAPOP), based on data from 24 Latin American countries, finds a positive correlation between beliefs that corruption is widespread among public officials and the likelihood of considering paying a bribe to be justified (Plata 2012).

In a broader reading, our article can be seen as a contribution to the literature on how social context affects individuals' willingness to engage in crime. Empirical evidence has shown that the decision to commit a crime is affected by the behavior of others. For instance, using data from U.S. cities and New York City neighborhoods, Glaeser, Sacerdote, and Scheinkman (1996) show that individuals are more likely to commit crimes when crime around them is widespread, especially less serious crimes. Using survey data from the United States, Sheffrin and Triest (1992) find that perceiving other taxpayers as dishonest increases the likelihood that an individual will evade taxes. Other studies have shown that neighborhood and peer effects are similarly important in explaining related behaviors such as academic cheating (Carrell, Malmstrom, and West 2008), fraud in emissions testing (Pierce and Snyder 2008), and shirking on the job (Ichino and Maggi 2000).

Finally, our article is part of a growing literature that uses experimental and quasi-experimental methods in order to understand how information about corruption shapes the behavior of citizens. Most of this work has concentrated on how information about corruption affects vote choice and political participation. For example, taking advantage of a natural experiment generated in Brazil by the randomized federal auditing of local governments, Ferraz and Finan (2008) show that mayors revealed to be corrupt lose electoral support. Focusing on the case of Mexico, Chong et al. (2015) found that distributing information about a corrupt incumbent decreases incumbent support as well as turnout. Winters and Weitz-Shapiro (2013) show that information about corruption decreases support for a hypothetical corrupt politician in Brazil, even when said politician performs well in office. This article extends upon such work by explicitly considering the role that citizens—as opposed to politicians or other officeholders—may play in actively propagating corruption throughout their societies. Moreover, our article is unique in the literature in that it provides microlevel experimental evidence on how "bad news" about corruption may lead citizens to perpetuate a vicious behavioral circle.

Measuring Citizens' Willingness to Engage in Corruption: The Joint Response Model

Accurately measuring whether a citizen would be willing to bribe (or has done so in the past) has long been recognized as one of the great challenges of empirical scholarship on corruption (e.g., Treisman 2007). Recognizing both the potential of social surveys to study corruption as well as the biases they invite when applied in standard form to sensitive issues, a number of scholars have begun to employ sensitive survey techniques (SSTs) in studies of this topic (Gingerich 2010, 2013; Malesky, Gueorguiev, and Jensen 2015). Following this lead, we utilize an SST-based approach in the current article. However, we do so in a novel way, by utilizing individual responses about corruption based on *both* a specific SST as well as upon direct questioning. We refer to the statistical framework

we utilize to analyze the protected and direct responses as the *joint response model*. In a recent contribution, Gingerich et al. (2015) show that utilizing a joint response approach provides all of the bias-reducing advantages of pure SST questioning while greatly enhancing the precision of parameter estimates.

Survey Questioning Format

The questioning strategy we utilized is easy to describe. First, survey respondents were presented with a question about willingness to bribe a police officer in order to avoid a traffic ticket using a particular SST format called the *crosswise model* (Tan, Tian, and Tang 2009). This technique provides anonymity to respondents via the commingling of responses about a sensitive issue (corruption) with responses about an innocuous question.[6] Next, at a later stage of the survey, respondents were queried directly about willingness to bribe a police officer (in the exact same context), with the option of "choose not to respond directly" provided to them. Observed responses about willingness to bribe were thus a discrete combination of responses under the protection afforded by the SST and the absence of protection under direct questioning.

Figure 2 presents the question about willingness to bribe based on the crosswise model. The respondent was presented with two statements and asked how many were true. The first statement states that the respondent's mother was born in October, November, or December. One can conceptualize affirmative responses to this statement as indicating membership in a nonsensitive group. The second statement, the one of interest, denotes a willingness to pay a bribe. The privacy of the respondent was protected by constraining the manner in which she was allowed to respond. There are only two potential responses: "A," indicating that either both statements are true *or* neither statement is true, and "B," indicating that only one of the two statements is true (but not specifying which is true). Since neither of the two responses necessarily indicates willingness to bribe, the respondent's anonymity is guaranteed.

In using the crosswise model, it is important to note that membership in the nonsensitive group is special in that it (1) must be known to each respondent but unknown to survey administrators (and known by each respondent to be unknown to administrators), (2) must be statistically independent of the sensitive trait of interest (willingness to bribe), (3) must have a proportion in the population of interest that is known in advance, and (4) must have a proportion that is different from 1/2 (otherwise, the crosswise responses would provide no information). Using the birth month of one's mother, as we did, helps ensure that nearly all respondents would know their own group assignment and that they would also be aware that the enumerator did not know their group assignment. Moreover, there is no plausible mechanism by which the birth month of one's mother should be tied to willingness to bribe, so the group indicator and the sensitive item are surely independent of one another. Finally, the population proportion of individuals belonging to the nonsensitive group is verifiable based on census records, meaning that such a group can be easily chosen such that the probability of membership differs arbitrarily from 1/2.[7]

The direct question about willingness to bribe, presented to respondents at the very end of the survey, asked them to respond only to the second statement presented in Figure 2. In this case, response options were "True," "False," and "I prefer not to respond."[8]

In order to calculate the probability of having one's mother born in the indicated interval of months, we conducted a nationally representative telephone survey of 1,200 Costa Ricans during July 2013. The survey queried respondents directly about the birthday of their mother and father. As an accuracy test, these responses were checked against data from Costa Rica's National Institute for Statistics and Censuses (INEC) on month of birth for newborns for the 2000–11 period. The figures from the phone survey self-reports and census data were essentially identical.[9]

Statistical Framework

Following the discussion above, we consider a setting in which each respondent i in a randomly selected sample of size n is first queried about her (unobservable) willingness to bribe, $\theta_i \in \{0$ ("unwilling"), 1 ("willing")$\}$ using the crosswise method and then later asked the same question directly. The (observable) combined response of respondent i to the two questions is denoted by the vector $Y_i =$

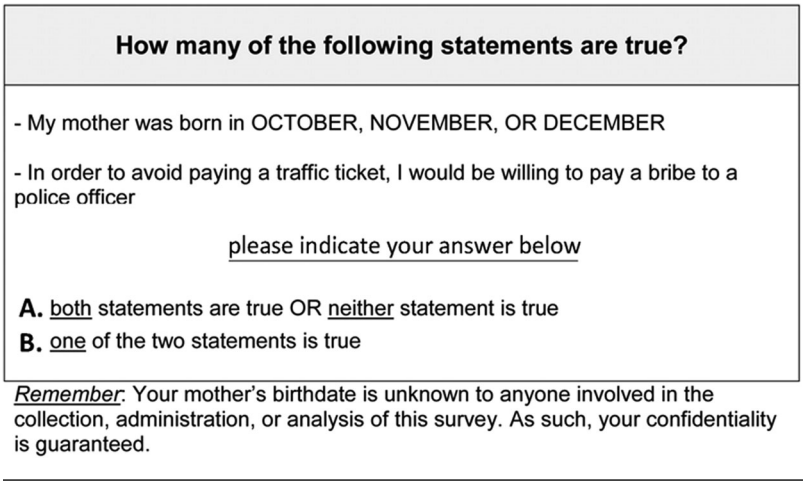
[6]The crosswise model is mathematically identical to the Warner version of the well-known randomized response technique, but it is administered without the use of a randomizing device.

[7]Recent studies have reported very good performance with the crosswise model in applications ranging from cheating by undergraduates to tax evasion (Jann, Jerke, and Krumpal 2012; Körndorfer, Krumpal, and Schmukle 2014; Kundt 2014).

[8]Detailed information about the enumerator scripts is provided in the supporting information.

[9]Appendix Table 4 in the supporting information compares the proportion of birthdays falling in the indicated months from the survey self-reports to the actual proportions for newborn births produced by INEC.

FIGURE 2 Crosswise Survey Item on Willingness to Bribe



(y_i^D, y_i^A), where $y_i^D = \{0$ ("False"), 1 ("True"), $\varnothing$ ("unwilling to respond directly")$\}$ is the observed response when i is asked to respond directly and $y_i^A \in \{0$ ("B"), 1 ("A")$\}$ is the observed ("anonymous") response when i is queried about bribery using the crosswise model. The observed response set is thus an array with six distinct elements, $\mathcal{Y} = \{(0, 0), (0, 1), (1, 0), (1, 1), (\varnothing, 0), (\varnothing, 1)\}$, with $k \in \mathcal{Y}$ representing an arbitrary element in this set. Without loss of generality, we can relabel responses as $Y_i \in \mathcal{Y} = \{1, 2, \ldots, 5, 6\}$, where each natural number 1, .., 6 represents one of the six distinct response combinations. For the responses using the crosswise technique, $z \neq 1/2$ will denote the probability that the first statement is true (e.g., the probability that the respondent's mother was born in the indicated interval of months). This quantity is known prior to collecting the data. (For the question displayed in Figure 2, $z = 0.264$.)

Our primary interest in this article resides in estimating the parameters of a model of the conditional probability of being willing to bribe given a respondent's experiences and observed characteristics. Let $\pi_i \equiv \mathbb{P}(\theta_i = 1|\mathbf{X}_i) = (1 + \exp(-\mathbf{X}_i^\top \beta))^{-1}$ where $\mathbf{X}_i$ is a vector of background characteristics and/or a treatment assignment recorded in the social survey along with a constant and β is the parameter vector. Since β reflects the influence of the experiences or characteristics of a respondent on her willingness to bribe, we refer to the elements of this vector as the explanatory parameters of our statistical model.

Our statistical framework rests on two key assumptions. The first is called *honesty given protection*: Given the protection afforded by the crosswise model, all respondents are assumed to respond honestly and as prompted by the technique (cf. Blair and Imai 2012; Gingerich 2010). In other words, lying is assumed to occur *only* when respondents are prompted to respond directly about their willingness to bribe. The second assumption is called *one-sided lying*. It holds that individuals who do not bear the sensitive trait never falsely claim that they do. Rather, the set of potential liars is limited to those respondents who do bear the sensitive trait.

The statistical model parameterizes patterns of evasiveness under direct questioning. In particular, let λ_θ^T, λ_θ^L, and $1 - \lambda_\theta^T - \lambda_\theta^L$ denote the probability that, when queried directly, a respondent whose status is θ tells the truth about her willingness to bribe, lies about it, or refuses to answer the question, respectively. Formally, one-sided lying implies that $\lambda_0^L = 0$. The assumption reflects the presumed direction of social desirability bias in sensitive surveys. If concerns about social desirability make it difficult for respondents with a sensitive trait to publicly divulge their status, those same concerns should ensure that respondents without the sensitive trait would have no incentive to falsely state that they bear the trait. Since this second set of parameters captures potential biases in responses generated by direct questioning, we refer to these as the diagnostic parameters of our statistical model.

The probability that a given respondent exhibits each combination of responses in the observed response set is presented in Table 1. Each cell of the table expresses the probability of observing the particular response combination represented by that cell.

Let $I(.)$ be an indicator function equal to 1 if its argument is true, 0 otherwise; let $\mathbb{P}_Y(k|\mathbf{X}_i)$ be the probability that respondent i's observed joint response is in category k given her background characteristics, the model for observed responses (e.g., the probabilities presented in Table 1), and the model for the conditional probability of

TABLE 1 Probability for Observed Data under Assumptions of Honesty Given Protection and One-Sided Lying

Y_i	Outcome	Probability
1	$(y_i^D = 0, y_i^A = 0)$	$z\lambda_0^T(1-\pi_i) + (1-z)\lambda_1^L\pi_i$
2	$(y_i^D = 0, y_i^A = 1)$	$(1-z)\lambda_0^T(1-\pi_i) + z\lambda_1^L\pi_i$
3	$(y_i^D = 1, y_i^A = 0)$	$(1-z)\lambda_1^T\pi_i$
4	$(y_i^D = 1, y_i^A = 1)$	$z\lambda_1^T\pi_i$
5	$(y_i^D = \varnothing, y_i^A = 0)$	$z(1-\lambda_0^T)(1-\pi_i) + (1-z)(1-\lambda_1^T-\lambda_1^L)\pi_i$
6	$(y_i^D = \varnothing, y_i^A = 1)$	$(1-z)(1-\lambda_0^T)(1-\pi_i) + z(1-\lambda_1^T-\lambda_1^L)\pi_i$

being willing to bribe; and let $\xi = (\lambda_1^T, \lambda_1^L, \lambda_0^T, \beta)^\top$ be the vector of parameters to be estimated. The log-likelihood function for the parameters given the observed data is written as

$$\ln L(\xi|Y, \mathbf{X}) = \sum_{i=1}^{n}\sum_{k=1}^{6} I(Y_i = k)\ln \mathbb{P}_Y(k|\mathbf{X}_i) \quad (1)$$

Note that if one simply wishes to calculate the (unconditional) proportion of individuals willing to bribe, one can write $\pi_i = \pi = \mathbb{P}(\theta_i = 1)$. In this case, $\xi = (\pi, \lambda_1^T, \lambda_1^L, \lambda_0^T)^\top$ and the log-likelihood function simplifies to

$$\ln L(\xi|Y) = \sum_{k=1}^{6} n_k \ln \mathbb{P}_Y(k), \quad (2)$$

where $n_k = \sum_{i=1}^{n} I(Y_i = k)$ is the number of respondents exhibiting response category k.

We utilize the expectation maximization (EM) algorithm to obtain the maximum likelihood estimates (MLEs) of the parameters in our statistical model. The EM algorithm, typically applied in incomplete-data settings, is particularly apposite for the setting studied in this article due to the (partial) unobservability of our outcome of interest.

The Information Experiment

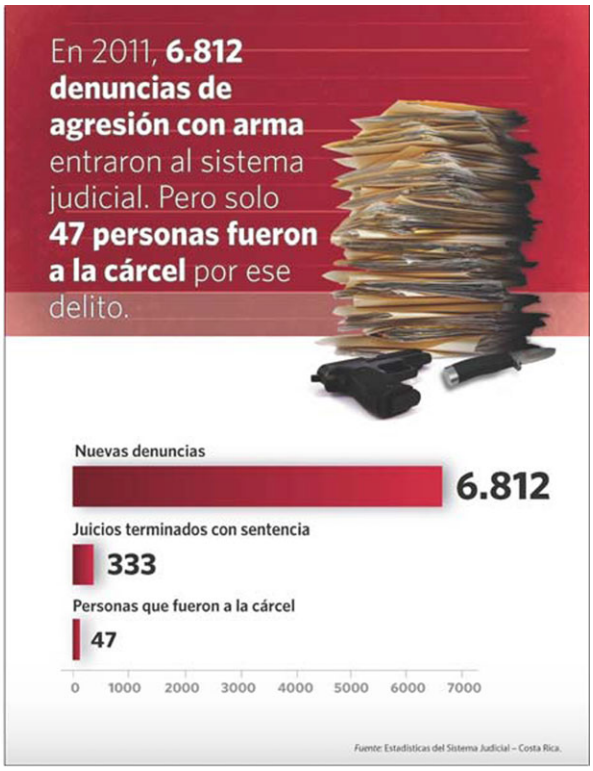
To study the effects of perceived corruption in society on an individual's willingness to engage in corrupt behavior, we combined an information experiment embedded in a household survey with the modeling framework developed above. The survey consisted of face-to-face interviews of 4,200 adults in the Gran Área Metropolitana (GAM), which includes 30 cantons in the provinces of Alajuela, Cartago, Heredia, and San José. The GAM is the principal urban center in Costa Rica. It contains approximately 2.6 million residents and accounts for 60% of the country's population. The survey was administered by the firm *Borge y Asociados* between October 2013 and April 2014.[10]

Rather than basing our analysis on the observed correlation between perceptions of corruption and willingness to engage in corrupt behavior, a strategy likely to suffer from potentially severe problems of confounding, we induce exogenous variation in beliefs about corruption via the random assignment of respondents to distinct informational treatments. Three informational treatments were employed: a corruption treatment, an inefficiency treatment, and a control condition. In the corruption treatment, respondents were presented with a flyer depicting the increasing percentage of Costa Ricans who have directly observed an act of corruption. A second treatment, the inefficiency treatment, was introduced as a placebo in order to test whether respondents were affected by the information included in the corruption treatment or by just the fact that they were given a flyer with negative information about the capacity of the Costa Rican state to deal with illicit behavior. In the inefficiency treatment, respondents were presented with a flyer presenting the (lack of) productivity of the legal system in dealing with a particular crime: assault with a deadly weapon. In the control condition, respondents were not presented with any flyer. Randomization of treatment assignment was programmed directly into the portable digital assistants (PDAs) the enumerators used to conduct the survey. Random assignment to different types of information ensured that, on average, groups of respondents were indistinguishable on both observable and unobservable characteristics. Appendix Table 5 in the supporting information provides evidence on balance in observable respondent characteristics across experimental groups.

The two flyers are presented in Figure 3. The flyer on the left-hand side is the corruption treatment. It states, "Did you know that corruption in Costa Rica has

[10]See the supporting information for more details on the survey methodology and execution.

FIGURE 3 The Two Treatments: Corruption and Judicial Inefficiency

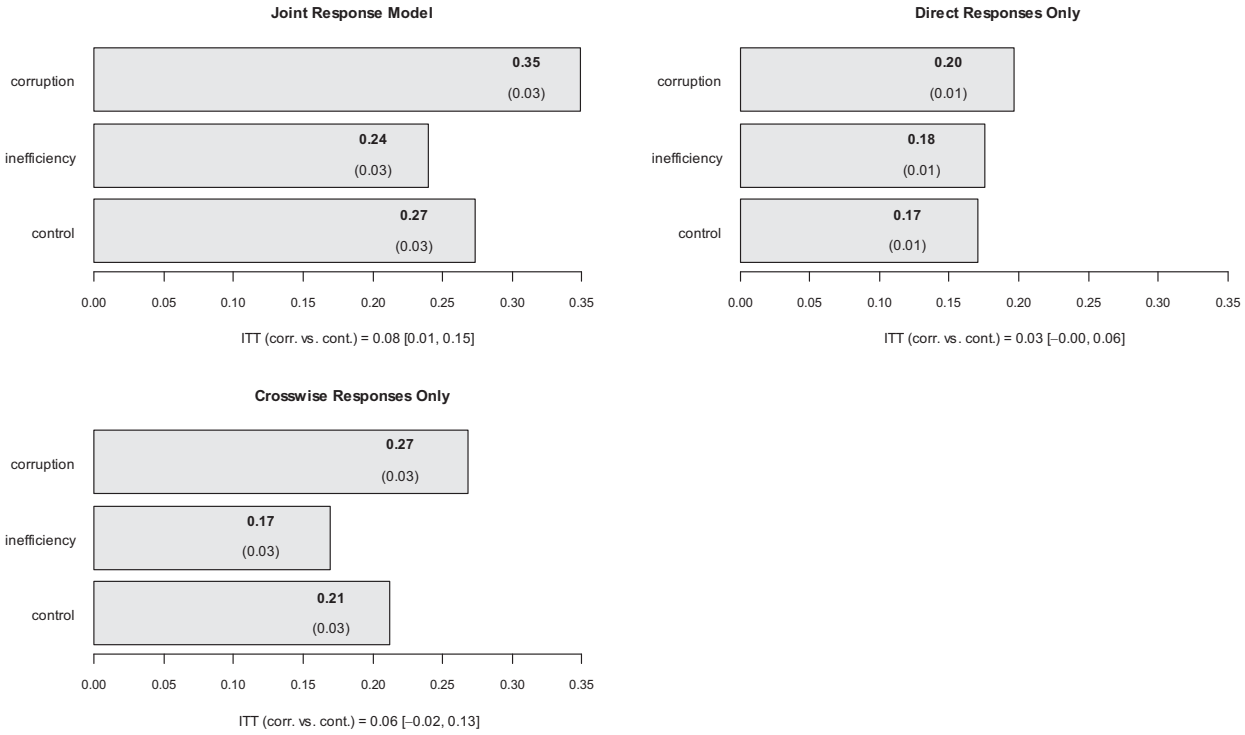


increased?" Below the statement appears a bar graph showing the increase in the percentage of Costa Ricans who had personally witnessed an act of corruption from 2006 (16%) to 2011 (24%). At the bottom right-hand side of the flyer, the source of the information, a nationally representative survey conducted by *Latinobarómetro*, is displayed. The flyer on the right-hand side is the inefficiency treatment. It states, "In 2011, 6,812 cases of assault with a deadly weapon entered into the judicial system. However, only 47 individuals were sent to jail for this crime." Below the statement appears a bar graph showing the relative magnitudes of the number of cases filed for this crime (6,812), the number of judicial decisions made on cases of the crime (333), and the number of individuals actually sent to jail (47). The source of the information, National Judicial Statistics, is displayed in the bottom right corner.

In the latter third of the survey (well after exposure to one of the three experimental conditions), respondents assigned to the two informational treatments were presented with a verification question that asked them to describe what the informational graphic they received was about. Subsequent to this, all respondents were prompted to respond to the aforementioned question about willingness to bribe, first in crosswise format and then, at the very end of the survey, in direct questioning format.

The purpose of the verification question was to identify individuals who were assigned to the information treatments but who failed to fully internalize the information they were given. We classify an individual as a full recipient of an information treatment if she was assigned to said treatment and could recall the basic content of the treatment. Respondents assigned to the corruption treatment were categorized as full recipients if they stated on the verification question that the informational graphic they received was about how "corruption has increased in recent years" or that it dealt with "something about corruption." According to this, 76% of respondents assigned to the corruption treatment were full recipients (1,065 out of 1,393). Respondents assigned to the judicial inefficiency treatment were categorized as full recipients if they stated on the verification question that the informational graphic they received was about how "there are many reports of crime but few people go to jail" or that it dealt with "something about how bad/inefficient the judicial system is." Sixty-five percent of respondents assigned to the judicial inefficiency treatment were full recipients according to this standard (904 out of 1,385).

FIGURE 4 Estimated Proportion of Respondents Willing to Bribe by Treatment Condition and Questioning Method



Note: Standard errors are in parentheses. The 95% confidence intervals for the average treatment effect are in brackets.

Corruption Does Corrupt

Intent-to-Treat Estimates

We begin our analysis by examining differences in the estimated proportion of respondents willing to bribe according to treatment assignment. These differences are the basis of our estimates of the intent-to-treat (ITT) effect: the average impact of random assignment to one of the three experimental conditions. Figure 4 presents the results utilizing our joint response approach, the direct survey responses, and the crosswise responses.

The results provide support for the self-fulfilling hypothesis that an individual's willingness to engage in corrupt behavior is affected by her exposure to information about the level of corruption in society. According to the estimates based on the joint response model, the proportion of respondents assigned to the corruption information treatment who would be willing to bribe a police officer was 0.35, whereas the proportion of respondents assigned to the control condition willing to do the same was only 0.27. Thus, the average effect of exposure to information about the increasing scope of corruption was equal to 0.08. This is a substantively large effect: Exposure to the corruption treatment was estimated to increase the proportion of respondents willing to bribe by 28%. In addition to being large in magnitude, the effect was statistically significant by conventional standards. Our placebo treatment—information about judicial inefficiency—did not have a statistically significant impact on willingness to bribe.

In addition to providing evidence that corruption corrupts, the figure also presents differences in response patterns across questioning techniques. In every treatment condition, estimates of willingness to bribe based on direct responses were below those based on the joint response model and crosswise responses only. Moreover, it would appear that desirability bias substantially attenuated downward the estimate of the effect of exposure to information about corruption: The impact estimate based only on direct responses was about half that based only on the crosswise responses and just about one-third as large as that based on the joint responses. The use of only the crosswise responses resulted in an estimated effect equal to 0.06, below but roughly similar to that based on the joint responses. However, this estimate was not statistically significant at conventional standards, due in part to the fact that it is based on an inefficient statistical

approach that failed to incorporate the potentially useful information available from direct responses.

The next step of our analysis consisted of the use of the joint response–modified logistic regression framework described earlier.[11] Table 2 depicts coefficient estimates showing the impact of the two information treatments on a respondent's willingness to offer a bribe to a police officer to avoid paying a traffic ticket. Two regression models were estimated: one in which the informational treatments were entered as the sole explanatory variables and one in which we included the age, gender, and education of the respondent. Previous findings in the literature suggest that women and older individuals might be less involved in corruption and/or less likely to condone corruption than men and younger individuals (Swamy et al. 2003; Torgler and Valev 2010) and that more educated individuals (or richer, the two variables often used as proxies) are more tolerant of corruption (Winters and Weitz-Shapiro 2013).

Both estimated models told a similar story: Exposing respondents to information about the growing scope of corruption in society made them significantly more likely to indicate a willingness to bribe a police officer to avoid paying a traffic ticket. Again, it thus appears that, as predicted by theory, corruption really does corrupt. The average effect of exposure to information about the growing scope of corruption was estimated to be 0.05 in the baseline model and 0.04 in the model with additional covariates. In both cases, the ITT estimates were statistically significant by conventional standards. As above, in neither of the estimations did the judicial inefficiency treatment have any discernible impact on willingness to bribe, suggesting that the effect of the corruption treatment was caused by the specific information contained in it and not by just the fact that respondents were exposed to some negative information about the capacity of the Costa Rican state to deal with illicit activity. In terms of background characteristics, men appeared to be substantially more inclined to bribe than women, younger respondents more inclined to bribe than older respondents, and individuals with incomplete secondary school education more inclined to bribe than individuals with some exposure to college (the baseline education category).[12]

[11] Unlike the joint response estimates presented in Figure 4, this approach pools the estimation of the diagnostic parameters across all three treatment conditions.

[12] As with any other information experiment, there are good reasons to believe that different types of citizens might react differently to the information provided. Individual characteristics such as gender, age, education, wealth, and prior beliefs about the level of corruption might interact with the information about rising levels of corruption. To study whether these characteristics have a conditioning effect on our experimental results, we estimated conditional intent-to-treat effects (CITTs) for various subgroups of the population based on these characteristics. None of the differences in the intent-to-treat estimates achieved statistical significance. See Tables 6–11 in the supporting information.

Local Average Treatment Effects

The estimates of the impact of information about corruption provided above understate the impact of actually internalizing such information. This is so because some of the respondents assigned to the information treatments failed to fully consume the information to such an extent that they could recall it accurately later. Since these individuals cannot be said to have received the treatment in a meaningful way, their presence in a given treatment group deflates the impact estimate associated with that treatment.

In order to estimate the causal impact of actually internalizing the information provided, we employ an instrumental variables approach. In our conceptualization, a respondent only receives an information treatment if she is assigned to it *and* has internalized the information in the treatment according to the criteria described earlier. Treatment assignment is thus an instrumental variable. Compliance with the instrument is perfect for respondents assigned to the control group (these individuals cannot internalize the information from a given information treatment because they have not received it) but assumed imperfect for respondents assigned to the two graphical displays. Thus, in the setting considered here, the instrumental variables estimator for a given information treatment is equal to the ITT divided by the compliance rate for that treatment (e.g., the rate of internalization). As is well known, this estimator estimates the local average treatment effect (Angrist, Imbens, and Rubin 1996): In our case, this is the average effect of receiving an information treatment for those individuals capable of internalizing the information contained therein.

Table 3 presents our estimates of the local average treatment effect associated with internalizing the information contained in the two treatments. These estimates are based upon employing the joint response model in order to estimate respondents' willingness to bribe. We present results based on estimating prevalence rates and diagnostic parameters separately for each treatment group as well as results based on our modified logistic regression framework (which pools the diagnostic parameters across treatment groups). In the case of the corruption treatment, the estimated local average treatment effects are substantively large and all statistically significant by conventional standards, supporting the notion

TABLE 2 Relationship between Information Treatments and Bribery of Police (Joint Response Model)

	Model 1			Model 2		
Parameters	**Estimate**	**SE**	**95% Int.**	**Estimate**	**SE**	**95% Int.**
Diagnostic parameters						
$\widehat{\lambda}_1^T$	0.61	0.03	[0.55, 0.67]	0.63	0.03	[0.58, 0.68]
$\widehat{\lambda}_1^L$	0.36	0.03	[0.30, 0.41]	0.34	0.03	[0.28, 0.39]
$\widehat{\lambda}_0^T$	0.97	0.00	[0.96, 0.98]	0.97	0.00	[0.96, 0.98]
Explanatory parameters						
Constant	−0.98	0.09	[−1.17, −0.80]	2.81	0.46	[1.97, 3.77]
Corruption treatment	**0.24**	**0.10**	**[0.04, 0.45]**	**0.23**	**0.11**	**[0.01, 0.43]**
Inefficiency treatment	**0.01**	**0.11**	**[−0.18, 0.22]**	**0.03**	**0.12**	**[−0.19, 0.24]**
Male	—	—	—	0.89	0.10	[0.70, 1.10]
Log (age)	—	—	—	−1.29	0.12	[−1.54, −1.06]
Education (base = some college)	—	—	—			
Primary or less	—	—	—	0.12	0.16	[−0.17, 0.40]
Secondary incomplete	—	—	—	0.40	0.15	[0.13, 0.68]
Secondary complete	—	—	—	0.29	0.15	[−0.01, 0.59]
Some technical	—	—	—	0.05	0.28	[−0.51, 0.62]
ITT (corruption vs. control)	**0.05**	**0.02**	**[0.01, 0.09]**	**0.04**	**0.02**	**[0.00, 0.08]**
		$n = 4{,}193$			$n = 4{,}192$	

TABLE 3 Local Average Treatment Effects of Internalizing Information Treatments (Outcome Measured Using Joint Response Model)

	LATE Estimates Diagnostic Parameters (Unpooled vs. Pooled)		
Information treatment	**Unpooled**	**Pooled: Model 1**	**Pooled: Model 2**
Corruption	0.10	0.07	0.05
	(0.05)	(0.03)	(0.03)
	[0.00, 0.20]	[0.00, 0.13]	[0.00, 0.11]
Inefficiency	−0.05	0.00	0.01
	(0.06)	(0.03)	(0.03)
	[−0.16, 0.05]	[−0.06, 0.07]	[−0.06, 0.07]

Note: Standard errors are in parentheses, and 95% confidence intervals are in brackets. Standard errors and confidence intervals were calculated via the nonparametric bootstrap.

that expectations about a widening societal scope of corruption drive willingness to bribe. The estimate of the local average treatment effect produced by the joint response approach without pooling diagnostic parameters was equal to 0.10; the modified logistic regression estimates based on pooling were equal to 0.07 and 0.05 (the former based on employing an explanatory model with no covariates and the latter based on a model with covariates). Since the vast majority of respondents assigned to the corruption treatment internalized the information therein (76%), the local average treatment effect estimates are only modestly larger in magnitude than the ITT estimates presented in the previous section. In terms of the judicial inefficiency treatment, we find that internalizing the information therein (65% of those assigned) had no statistically significant impact on willingness to bribe. Again, our conclusion is that citizens are responding specifically to the information about the scope of corruption and not to generically negative information about the capacity of the Costa Rican state.

Conclusion

This article provides the first set of experimental findings in favor of the proposition that corruption corrupts. Utilizing data from a household survey conducted in Costa Rica, one that combined both a survey experiment and a novel format for asking sensitive questions (the joint response model), we find that exposing citizens to information about the growing scope of corruption in their society made them individually more disposed to engage in corruption. In this way, our article offers empirical support to a large theoretical literature on corruption that has long claimed that the phenomenon is characterized by strategic complementarities. At a methodological level, our article shows how survey responses generated by sensitive survey techniques and direct questioning can be combined to study sensitive issues.

In the last decade, Costa Rica has witnessed a number of public scandals unusual for its political history, as well as rising levels of corruption. These factors might have made our respondents particularly susceptible to the information provided in our experiment. Could we expect similar results in other countries? As with any other information experiment, the information provided can only have an effect if the respondents find it credible. Thus, we would not expect similar results were our experiment conducted in countries with long-lived reputations for cleanliness in government. Nor would we expect a strong effect of information in the opposite direction (decreasing scope of corruption) in countries with equally long-lived reputations for widespread corruption. In both instances, citizens' priors about the scope of corruption would likely be quite difficult to dislodge. Countries where reputations for corruption appear to be in flux would seem to be the most likely candidates to exhibit informational effects of the kind reported in these pages.

In this respect, it is worth noting that Costa Rica's recent experience is not particularly unusual. According to the *Latinobarómetro*—the source we used in our experiment—two other countries (out of the 18 in the sample) have experienced recent increases in corruption. While the percentage of citizens observing an act of corruption in Costa Rica went from 16 to 24 between 2006 and 2011, this percentage increased from 12 to 14 in Colombia, and from 17 to 21 in the Dominican Republic during the same period. The case of Chile is also worth considering. Even more so than Costa Rica, Chile has long had a reputation for probity in government. However, the last several years have produced numerous scandals involving the illicit financing of campaigns as well as serious instances of tax fraud, money laundering, and influence peddling by a small, politically connected elite. In theory, scandals such as these could perturb beliefs in a direction conducive to the spread of corruption along the lines of what we found with our subjects in Costa Rica. Outside of the Americas, Spain, which has experienced an unprecedented wave of corruption scandals over the last decade (Lopez-Valcarcel, Jiménez, and Perdiguero 2014), would be another country for which the dynamics identified in this article may apply. The existence of these similar cases notwithstanding, an important task for future research is to sort out the degree to which the results encountered here extend to polities where corruption levels and expectations are relatively more stable than in Costa Rica.

In closing, we would like to underline some practical lessons to be taken from this article. First and foremost, we believe our findings encourage the adoption of additional nuance in policy-oriented discussions about the role of transparency in reducing corruption. Greater transparency is often heralded by those in the international development community as one of the most important antidotes for reducing corruption. A number of influential papers, cited earlier, suggest that providing voters with information on corruption can promote a virtuous response in which citizens attempt to vote corrupt politicians out of office. We do not doubt that transparency has an important role to play in the fight against corruption, especially at election time. However, the results we encountered in this study lead us to believe that transparency about corruption might be more of a double-edged sword than many have previously thought.

In particular, our findings lead us to be concerned about the potentially damaging effects of shaming campaigns that attempt to galvanize public opinion against corruption by widely disseminating the message that large swaths of public officials are on the take or that the accumulated losses of corruption are enormous. Examples of such campaigns include the "I paid a bribe" websites in India, Kenya, and Pakistan (which provide real-time information on anonymous, self-reported bribery payments); the installation of the so-called *abusometro*, an electronic billboard located in Mexico City that gives citizens a running tally of estimated public education funds lost to corruption; and the corruption bus tours that expose citizens to the sites of corrupt exchange and the fruits of corruption in cities such as Monterrey, Mexico, and Prague, Czech Republic. Although these efforts are well intentioned, our findings suggest that constantly nailing inconvenient facts into citizens' heads about the scope of corruption can shape expectations about the behavior of public officials in such a way as to perpetuate the very problem such campaigns are designed to solve.

At the end of the day, many forms of corruption are not simply a matter of what officials decide to do to citizens, but rather a matter of what officials and citizens conspire to do *together* against the interests of the public at large. As the saying goes, it takes two to tango, and the dance of corruption requires at least two willing partners inclined to risk their reputation, patrimony, and possibly their liberty in order to cement a corrupt exchange with an individual about whom they typically know fairly little. In such contexts, societal expectations about what typical citizens and officials are inclined to do will be central in delimiting the risk–reward calculus of both parties to the exchange. Transparency of a kind that conveys the message that there is little to be risked but much to be gained from pursuing such an exchange is a type of transparency that most polities are best left without.

References

Accinelli, Elvio, and Edgar J. Sánchez Carrera. 2012. "Corruption Driven by Imitative Behavior." *Economics Letters* 117(1): 84–87.

Açemoglu, Daron. 1995. "Reward Structures and the Allocation of Talent." *European Economic Review* 39(1): 17–33.

Aidt, Toke S. 2003. "Economic Analysis of Corruption: A Survey." *The Economic Journal* 113(491): F632–52.

Anderson, Christopher J., and Yuliya V. Tverdova. 2003. "Corruption, Political Allegiances and Attitudes towards Government in Contemporary Democracies." *American Journal of Political Science* 47(1): 91–109.

Anduiza, Eva, Gallego Aina, and Jordi Mu noz. 2013. "Turning a Blind Eye: Experimental Evidence of Partisan Bias in Attitudes toward Corruption." *Comparative Political Studies* 46(12): 1664–92.

Andvig, Jens C., and Karl O. Moene. 1990. "How Corruption May Corrupt." *Journal of Economic Behavior & Organization* 13(1): 63–76.

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91(434): 444–55.

Balafoutas, Loukas. 2011. "Public Beliefs and Corruption in a Repeated Psychological Game." *Journal of Economic Behavior & Organization* 78(1): 51–59.

Becker, Sascha O., Peter H. Egger, and Tobias Seidel. 2009. "Common Political Culture: Evidence on Regional Corruption Contagion." *European Journal of Political Economy* 25(3): 300–10.

Blair, Graeme, and Kosuke Imai. 2012. "Statistical Analysis of List Experiments." *Political Analysis* 20(1): 47–77.

Cadot, Oliver. 1987. "Corruption as a Gamble." *Journal of Public Economics* 33(2): 223–44.

Cameron, Lisa, Ananish 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): 843–51.

Carrell, Scott E., Frederick V. Malmstrom, and James E. West. 2008. "Peer Effects in Academic Cheating." *Journal of Human Resources* 43(1): 173–207.

Caselli, Francesco, and Massimo Morelli. 2004. "Bad Politicians." *Journal of Public Economics* 88(3): 759–82.

Chong, Alberto, Ana De La O, Dean Karlan, and Leonard Wantchekon. 2015. "Does Corruption Information Inspire the Fight or Quash the Hope? A Field Experiment in Mexico on Voter Turnout, Choice, and Party Identification." *Journal of Politics* 77(1): 55–71.

Dal Bó, Ernesto, Pedro Dal Bó, and Rafael Di Tella. 2006. "'Plata o Plomo?': Bribe and Punishment in a Theory of Political Influence." *American Political Science Review* 100(1): 41–53.

Dollar, David, Raymond Fisman, and Roberta Gatti. 2001. "Are Women Really the "Fairer" Sex? Corruption and Women in Government." *Journal of Economic Behavior & Organization* 46(4): 423–29.

Dong, Bin, Uwe Dulleck, and Benno Torgler. 2012. "Conditional Corruption." *Journal of Economic Psychology* 33(3): 609–27.

Dong, Bin, and Benno Torgler. 2012. "Corruption and Social Interaction: Evidence from China." *Journal of Policy Modeling* 34(6): 932–47.

Esarey, Justin, and Gina Chirillo. 2013. "'Fairer Sex' or Purity Myth? Corruption, Gender, and Institutional Context." *Politics and Gender* 9(4): 390–413.

Ferraz, Claudio, and Federico Finan. 2008. "Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes." *Quarterly Journal of Economics* 123(2): 703–45.

Fisman, Raymond, and Edward Miguel. 2007. "Corruption, Norms, and Legal Enforcement: Evidence from Diplomatic Parking Tickets." *Journal of Political Economy* 115(6): 1020–48.

Gingerich, Daniel W. 2009. "Ballot Structure, Political Corruption, and the Performance of Proportional Representation." *Journal of Theoretical Politics* 21(4): 509–41.

Gingerich, Daniel W. 2010. "Understanding Off-the-Books Politics: Conducting Inference on the Determinants of Sensitive Behavior with Randomized Response Surveys." *Political Analysis* 18(3): 349–80.

Gingerich, Daniel W. 2013. *Political Institutions and Party-Directed Corruption in South America: Stealing for the Team.* Cambridge: Cambridge University Press.

Gingerich, Daniel W., Virginia Oliveros, Ana Corbacho, and Mauricio Ruiz-Vega. 2015. "When to Protect? Using the Crosswise Model to Integrate Protected and Direct Responses in Surveys of Sensitive Behavior." *Political Analysis* Forthcoming. doi:10.1093/pan/mpv034

Glaeser, Edward L., Bruce Sacerdote, and Jose A. Scheinkman. 1996. "Crime and Social Interactions." *Quarterly Journal of Economics* 111(2): 507–48.

Goel, Rajeev K., and Michael A. Nelson. 2007. "Are Corrupt Acts Contagious? Evidence from the United States." *Journal of Policy Modeling* 29(6): 839–50.

Goel, Rajeev K., and James W. Saunoris. 2014. "Global Corruption and the Shadow Economy: Spatial Aspects." *Public Choice* 161(1): 119–39.

Hauk, Esther, and Maria Saez-Marti. 2002. "On the Cultural Transmission of Corruption." *Journal of Economic Theory* 107(2): 311–35.

Ichino, Andrea, and Giovanni Maggi. 2000. "Work Environment and Individual Background: Explaining Regional Shirking Differentials in a Large Italian Firm." *Quarterly Journal of Economics* 115(3): 1057–90.

Jann, Ben, Julia Jerke, and Ivar Krumpal. 2012. "Asking Sensitive Questions Using the Crosswise Model: An Experimental Survey Measuring Plagiarism." *Public Opinion Quarterly* 76(1): 32–49.

Korndörfer, Martin, Ivar Krumpal, and Stefan C. Schmukle. 2014. "Measuring and Explaining Tax Evasion: Improving Self-Reports Using the Crosswise Model." *Journal of Economic Psychology* 45(December): 18–32.

Kundt, Thorben Christian. 2014. "Applying 'Benford's Law' to the Crosswise Model: Findings from an Online Survey on Tax Evasion." Unpublished manuscript. http://ssrn.com/abstract=2487069.

Lopez-Valcarcel, Beatriz G., Juan Luis Jiménez, and Jordi Perdiguero. 2014. "Danger: Local Corruption Is Contagious!" Unpublished manuscript. http://www.alde.es/encuentros/anteriores/xviieea/trabajos/g/pdf/179.pdf.

Lui, Francis T. 1986. "A Dynamic Model of Corruption Deterrence." *Journal of Public Economics* 31(2): 215–36.

Malesky, Edmund J., Dimitar D. Gueorguiev, and Nathan M. Jensen. 2015. "Monopoly Money: Foreign Investment and Bribery in Vietnam: A Survey Experiment." *American Journal of Political Science* 59(2): 419–39.

Márquez, Miguel A., Javier Salinas-Jiménez, and Ma del Mar Salinas-Jiménez. 2011. "Exploring Differences in Corruption: The Role of Neighboring Countries." *Journal of Economic Policy Reform* 14(1): 11–19.

Medina, Luis Fernando. 2007. *A Unified Theory of Collective Action and Social Change*. Ann Arbor: University of Michigan Press.

Mishra, Ajit. (2006). "Persistence of Corruption: Some Theoretical Perspectives." *World Development* 34(2): 349–58.

Pierce, Lamar, and Jason Snyder. 2008. "Ethical Spillovers in Firms: Evidence from Vehicle Emissions Testing." *Management Science* 54(11): 1891–1903.

Plata, Juan Camilo. 2012. "To Bribe or Not to Bribe." *AmericasBarometer Insights* 72: 1–8.

Ryvkin, Dmitry, and Danila Serra. 2012. "How Corruptible Are You? Bribery under Uncertainty." *Journal of Economic Behavior & Organization* 81(2): 466–77.

Sheffrin, Steven M., and Robert K. Triest. 1992. "Can Brute Force Deterrence Backfire? Perceptions and Attitudes in Taxpayer Compliance." In *Why People Pay Taxes: Tax Compliance and Enforcement*, ed. J. Slemrod. Ann Arbor: University of Michigan Press, 193–218.

Sung, Hung-En. 2003. "Fairer Sex or Fairer System? Gender and Corruption Revisited." *Social Forces* 82(2): 703–23.

Swamy, Anand, Stephen Knack, Young Lee, and Omar Azfar. 2003. "Gender and Corruption. " In *Democracy, Governance and Growth*, ed. Stephen Knack. Ann Arbor: The University of Michigan Press, 191–224.

Tan, Ming T., Guo-Liang Tian, and Man-Lai Tang. 2009. "Sample Surveys with Sensitive Questions: A Nonrandomized Response Approach." *The American Statistician* 63(1): 9–16.

Tirole, Jean. 1996. "A Theory of Collective Reputations (with Applications to the Persistence of Corruption and to Firm Quality)." *Review of Economic Studies* 63(1): 1–22.

Torgler, Benno, and Neven T. Valev. 2010. "Gender and Public Attitudes toward Corruption and Tax Evasion." *Contemporary Economic Policy* 28(4): 554–68.

Treisman, Daniel. 2007. "What Have We Learned about the Causes of Corruption from Ten Years of Cross-National Empirical Research?" *Annual Review of Political Science* 10: 211–44.

Winters, Matt S., and Rebecca Weitz-Shapiro. 2013. "Lacking Information or Condoning Corruption: When Do Voters Support Corrupt Politicians?" *Comparative Politics* 45(4): 418–36.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

- The Corrupting Influence of Expectations about Corruption: A Formal Model
- Proofs for the Formal Model
- Survey Methodology
- Focus Groups
- Phone Survey
- Enumerator Scripts