UNITED NATIONS UNIVERSITY
UNU-WIDER

WIDER Working Paper 2016/137

Deterrence, peer effect, and legitimacy in anti-corruption policy-making

An experimental analysis

Amadou Boly,[1] Robert Gillanders,[2] and Topi Miettinen[2]

November 2016

Abstract: In our framed laboratory experiment, two Public Officials, A and B, make consecutive decisions regarding embezzlement from separate funds. Official B observes Official A's decision before making their own. There are four treatments: three with deterrence and one without. We find a peer effect in embezzlement in that facing an honest Official A reduces embezzlement by Official B. Likewise, deterrence matters in that higher detection probabilities significantly decrease embezzlement. Crucially, detection is more effective in curbing embezzlement when chosen by an honest Official A compared to a corrupt Official A at almost all individual detection levels. This 'legitimacy' effect may help explain why anti-corruption policies can fail in countries where the government itself is believed to be corrupt.

Keywords: corruption, deterrence, embezzlement, laboratory experiment, legitimacy, peer effect
JEL classification: C91, D03, D73, K42

Acknowledgements: This research project started while Amadou Boly was a Research Fellow at UNU-WIDER. He acknowledges support from UNU-WIDER project 'Macro-economic Management (M-eM)'. The views expressed here are those of the authors and do not necessarily represent or reflect those of the African Development Bank.

We thank the Busara Center for Behavioral Economics for conducting the experiments. We would also like to thank participants at the African Development Bank internal seminar series, the 2016 Canadian Economic Association Annual Conference, the 2016 PEGNet Conference, and the UNU-MERIT/MGSoG seminar series for comments and suggestions.

---

[1]African Development Bank (AfDB), Development Research Department, Abidjan, Cote d'Ivoire, corresponding author: a.boly@afdb.org; [2]Department of Economics, Hanken School of Economics, Helsinki, Finland.

This study has been prepared within the UNU-WIDER project on 'Macro-economic management (M-eM)'.

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9256-181-9

Typescript prepared by Lesley Ellen.

The United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Denmark, Finland, Sweden, and the United Kingdom.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

1 Introduction

Corruption has been found to have undesirable effects on key economic metrics such as macroeconomic growth (Mauro 1995), firm growth (Fisman and Svensson 2007), and income equality and poverty (Gupta et al. 2002). On account of these large costs of corruption and the fact that corruption is particularly prevalent in developing and transition economies, anti-corruption laws and policies often constitute important elements in both internally and externally initiated reforms and development plans. A commonly advocated anti-corruption approach is deterrence, which is justified by appeals to models of rational criminal behaviour. These models assume that an illegal act, such as corruption, is preferred and chosen if its net expected benefit is higher than that of legal alternatives (Becker 1968; Paternoster 1987; Eide et al. 2006). As a result, government authorities can increase compliance with the law by increasing the risks (probability of detection) and/or costs (severity of sanctions) associated with corrupt transactions.

The available experimental evidence from Kenya suggests that the existence of a probability of detection and punishment can indeed curb corruption (for example: Abbink et al. 2002; Schulze and Frank 2003; Olken 2007; Hanna et al. 2011). In a typical corruption experiment, a control treatment with zero detection probability is compared to an experimental treatment with one positive level of detection probability exogenously imposed by the researcher (see e.g. Abbink et al. 2002; Serra 2012). The present paper builds on this literature by considering environments in which the level of the detection is endogenously chosen, allowing us to analyse the role of legitimacy and peer effect in anti-corruption policy-making.[1]

Specifically, in the first treatment (called NoD for No Detection) of our framed corruption experiment, there are two Public Officials, called A and B. They both receive a salary and are entrusted with separate funds to be spent on social projects. Each public official has the opportunity to embezzle from the fund under his/her control before sending the remaining amount to a recipient. The recipient is different for each public official and is chosen randomly from a list of local NGOs and charities. Embezzlement is inefficient from a social point of view as the amount sent to each recipient is doubled while the amount embezzled is not. Three additional treatments incorporate a mechanism for detection and punishment wherein detection automatically leads to punishment, which entails a loss of all earnings for the period. In the Endogenous and Discretionary (ED) treatment, Public Official A must choose a level of detection probability which applies only to Public Official B. In the Endogenous and Non-Discretionary (END) treatment, Public Official A has the same power over the anti-corruption policy but detection and punishment applies to both public officials. The Exogenous and Non-Discretionary (XND) treatment sees both public officials face an exogenous probability of detection of 30 per cent.

Authorities are considered legitimate when the public views them as having both the legal and the moral authority for law enforcement (Tyler 2006). Legitimacy enhances compliance with the law even when the likelihood of sanctions is low (Tyler 2006). In contrast, a lack of legitimacy could translate into behaviour contrary to that sought, resulting in non-compliance with the law or even increased criminal behaviour (Kagan and Scholz 1984; Fehr and Rockenbach 2003). We operationalize this concept in our experimental design by allowing the 'public official' who

[1] See e.g. Fehr and Rockenbach (2003) and Falk and Kosfeld (2006) for labour market experiments with endogenous monitoring.

chooses the strength of the detection probability to be corrupt. These decisions are observed by a second public official who then makes his or her own decision regarding corruption.

Our findings suggest that there is a peer effect in corruption whereby a corrupt policy maker generates corruption in others. We also confirm the deterrent effects of detection probability in that increasing the detection probability significantly reduces the likelihood and level of embezzlement. However, this raw deterrent effect of monitoring and punishment is only present in an institutional setting in which the policy maker is exempt from its provisions, namely the ED treatment. In a setting with equality before the law (END treatment), the effectiveness of monitoring and punishment on the behaviour of the second public official is found to depend on the legitimacy of the policy maker. Specifically, *ceteris paribus*, we find that in this institutional setting, monitoring and punishment have an effect on the corrupt behaviour of the other participant only when the policy maker is honest. In other words, a lack of legitimacy undermines the effectiveness of deterrence as an anti-corruption mechanism. This legitimacy effect is not present in the setting with procedural asymmetry (ED treatment). Finally, anti-corruption measures can be put in place by national authorities or possibly promoted by external actors and our design also allows us to explore whether the endogeneity of anti-corruption measures matters by focusing on the highest and lowest levels of detection. For the highest level of detection, the results suggest that endogeneity matters only when a policy maker is honest and equality before the law prevails. In this case, an externally imposed anti-corruption policy may be less effective in lowering the likelihood of embezzlement than a policy selected by an honest policy maker who is subject to the anti-corruption mechanism herself. We also find that a zero level of detection results in higher likelihood and level of embezzlement when it is endogenously chosen rather than exogenously set.

Recent experimental literature has pointed to a beneficial effect of leadership in public good provision (Moxnes and Van der Heijden 2003; Güth et al. 2007). Specifically, the literature has shown that the benefits of leadership are lost if there are pronounced differences in the endowments, economic incentives, or information among the participants (Levati et al. 2007; Cappelen et al. 2015). Our findings complement this literature by showing that leadership in the provision of a public good (charity contribution) does not trigger significantly higher contributions by the followers where deterrence institutions apply discriminatorily to two parties. When the setting is symmetric, however, there are strong positive effects on the contributions of the followers and more so if strong deterrence institutions are applied for both the leader and the follower. This also illustrates how legitimacy and procedural justice interact in the effectiveness of anti-corruption policies.

We contribute to the literature on the effects of deterrence on corruption in several ways. First, to the best of our knowledge, this is the first experiment to demonstrate that a non-monetary factor such as legitimacy can affect the effectiveness of anti-corruption monitoring and punishment in a significant way, to the extent that the same policy actions produce different outcomes. As deterrence is more effective when chosen by honest officials, legitimacy can reduce the costs of enforcement. This is in line with Levi and Sacks (2009) who argue that citizens who perceive a regime as legitimate are more likely to comply with its precepts (even when the probability that non-compliance would be detected is low). Also, our results may help explain why anti-corruption policies can fail in countries where the government itself is believed (or known) to be corrupt. The governments of many developing and transition economies clearly face such legitimacy concerns. It can be noted that the legitimacy effect is only present when the institutional framework features procedural symmetry (equality before the law) in terms of who the anti-corruption mechanism applies to. The implications of our findings may extend to other policy-making situations where the actions of the policy maker run contrary to the objective of the promoted policy. For example, a policy maker trying to curb tax evasion while putting his/her own wealth in a fiscal paradise may face legitimacy issues that will affect the effectiveness of the said policy. Finally, as there is typically no significant difference between endogenously and exogenously chosen detection levels at a high enough level, our findings suggest that monitoring (if well implemented) can be as effective whether decided upon by a national authority or externally imposed by international partners. Furthermore, when the domestic policy maker is himself corrupt and subject to the provisions of his own policies, externally imposed monitoring is more effective at deterring corruption than an equally strong internally chosen policy, likely due to a legitimacy issue. Overall, the implication of our findings should be of interest and practical value to anti-corruption advocates and policy makers.

This paper is organized as follows. Section 2 discusses the relevant literature including the existing experimental literature on the effects of deterrence on corruption. In Section 3, we describe the main features of experimental design. Section 4 presents our results, which give rise to a theoretical treatment in Section 5. We conclude in Section 6.

2 Literature review and further motivation

As mentioned above, the potential of monitoring and punishment as an anti-corruption tool has been well studied in the experimental literature. The ground-breaking contribution in this regard can be found in Abbink et al. (2002). Abbink et al. find that even a very small exogenously determined probability of being caught coupled with a severe punishment can significantly and meaningfully reduce the likelihood of engaging in bribery. The effectiveness of this type of anti-corruption policy is also evident in the complex experimental setting used in Azfar and Nelson (2007) and Barr et al. (2009). Evidence from the field is provided by Olken (2007) who finds that government audits are effective at reducing corruption in the context of Indonesian infrastructure projects.

However, some studies have reached somewhat different conclusions. Schulze and Frank (2003) conclude that monitoring and punishment damages intrinsic motivation. Their experiment has an exogenous probability of detection that increases with the bribe taken. In this context, they find that monitoring reduces the number of subjects that choose the highest level of bribe but the average bribe actually increases. Serra (2012) finds that while low-level monitoring does not deter corruption alone, it is effective in a mixed top-down and bottom-up accountability system. Overall, these studies suggest that monitoring and punishment can be at least an important element of an effective anti-corruption strategy. We demonstrate that this effect holds in our experimental framework.

We then move on to show that this effectiveness can be mitigated by the legitimacy (here being corrupt or honest) of the person enacting the policy. While long and widely studied outside economics (see e.g. Weber 1964; Kornhauser 1984; Tyler 1990; or Papachristos et al. 2012), legitimacy has received attention only more recently in economics with a few papers underscoring its relevance theoretically (see e.g. Schnellenbach 2007; Basu 2015; Akerlof 2016) and empirically (Chen 2013). Akerlof (2016) explores the constraints that the need for legitimacy imposes on organizational behaviour outcomes such as the rejection of overqualified workers or above-market-clearing wages. Using a dataset on World War I deserters, Chen (2013) finds limited evidence that deserters' executions by the British army deterred absences. In contrast, the higher execution rate of Irish soldiers compared to British soldiers, regardless of the crime, stimulated absences, particularly Irish absences. We contribute to the literature on legitimacy by presenting evidence that an anti-corruption policy maker's decision to act corruptly reduces the effectiveness of any given level of monitoring and punishment chosen by the policy maker. This finding could reflect a process in which he is delegitimized in the eyes of others who are subject to the provisions of his policy. Tyler (2006) suggests that a sense of legitimacy helps people to voluntarily obey the law, therefore implying that a loss of legitimacy could lead to a decrease in compliance with the law.

Several experimental studies show that others' behaviour can influence an individual's own attitudes and behaviour. For example, it has been found that most people contribute more to public goods if others do so (Brandts and Schram 2001; Bardsley and Sausgruber 2005; Fischbacher and Gächter 2010) or that tax compliance depends on the behaviour of others in society (Fortin et al. 2007; Lefebvre et al. 2015).[2] In particular, d'Adda et al. (2014) run an experiment that allows for behaviour that is reminiscent of corruption and show that groups with (likely) dishonest leaders are more likely to cheat. Beams et al. (2003) find that the declared willingness of the subjects in their sample of accounting students to engage in insider trading increases with the perceived unfairness of the laws that bars such trading. Jones and Kavanagh (1996) conclude that the ethical behaviour of employees is influenced by the ethical behaviour of their peers and managers. Similarly, Dineen et al. (2006) conclude that the behavioural integrity of supervisors modifies the effect of their guidance on employee behaviour. Their results show that guidance improves organizational citizenship and reduces deviant behaviour when the supervisors are perceived to have integrity. However, when they are perceived to be lacking in this regard, their guidance is harmful in terms of desirable employee behaviour. Pierce and Snyder (2008) demonstrate that a firm's ethical norms can influence those of its workers. Specifically, they find that the pass rate of vehicle inspectors adjusts to conform to the norm prevailing at the facility in which they are working. The possibility that the corrupt behaviour of others can, from a social standpoint, negatively influence an individual's own attitudes and behaviour regarding corruption has also been studied non-experimentally with the results of Gatti et al. (2003) and Dong et al. (2012) suggesting that this is indeed the case. Our experiment provides additional evidence by allowing for the first-moving policy maker to 'set the tone' of the organization that the subjects find themselves operating in. Indeed, Lambsdorff (2015) argues that the 'tone at the top' is '… [m]aybe the most important factor in fighting corruption…' (Lambsdorff 2015: 10).

In sum, these literatures argue for and point to a clear role for peer effects, deterrence, legitimacy, and organizational design (i.e. institutions) in the determination of individual behaviour. Our design builds on this existing work by examining the effects of these factors on corruption.

3 Experimental design

The data used in this paper were generated from a framed laboratory experiment which was carried out at the Busara Center for Behavioral Economics, in Nairobi, Kenya. Our subjects are mostly university students from the University of Nairobi and they come from a variety of disciplines. In the remainder of this section, we outline our basic procedure and describe our experimental treatments in detail.

---

[2] These peer effects extend to various areas such as donations (Shang and Croson 2009; Smith et al. 2015), academic achievements (Sacerdote 2011), work effort supply (Falk and Ichino 2006; Bandiera et al. 2010), and other legally deviant and norm-breaking behaviours (Bikhchandani et al. 1998; Keizer et al. 2008).

3.1 Procedure

Busara staff members read instructions out loud at the start of each session. After this, subjects were invited to ask any questions that they might have. Their understanding of the task at hand was then tested with comprehension questions. The duration of each session was roughly one hour.

Our framed laboratory experiment mirrors a situation in which public officials have the opportunity to embezzle public funds. Embezzlement refers to a situation in which a corrupt actor misuses another party's resources to his own (direct or indirect) benefit. Crucially, the corrupt actor has legal access to the resources but not legal ownership. We model this in the following way. Our participants take one of two roles, Public Official A or Public Official B, which they keep throughout the experiment and play a sequential move game. New pairs consisting of one of each type of public official are formed randomly at the start of each round. Payoffs are expressed in terms of Experimental Currency Units (ECU) during the sessions before being converted to Kenyan Shillings at the end of the experiment at a rate of 8ECU to 1Ksh.

Public Official A and Public Official B are each paid a salary of 1,140ECU at the start of every round. They are then each allocated a fund amounting to 2,280ECU, which they are aware is intended to be spent on 'social projects'. Public Official A moves first and has to choose whether to keep 0ECU or 760ECU from the social fund. If Public Official A chooses to keep 760ECU, this amount is added to his payoff for the round. The balance (2,280—Amount Kept) is multiplied by 2 and, after conversion into Kenyan Shillings, is sent to a recipient, called Recipient 1. This recipient is randomly selected at the end of the experiment from a list of local non-governmental organizations (NGOs) and local charity funds. Carrying out our experiment in Kenya and using real donations to local NGOs adds further ecological validity to our study.[3]

After observing the choice of Public Official A, Public Official B makes his/her decisions. Whereas Public Official A faced a binary embezzlement choice, Public Official B can opt to embezzle any whole number between 0 and 2,280 from the social funds under his control. The amount that Public Official B chooses to keep is transferred to his/her private account. As was the case with the funds passed on by Public Official A, the remainder of the fund (2,280—Amount Kept) is doubled, converted into Kenyan Shillings, and sent to a recipient. However, this recipient, called Recipient 2, is different from Recipient 1 and is also randomly selected from a list of local NGOs and local charity funds.[4]

Each session lasted for 40 independent rounds. If there were 22 participants in the laboratory, then 11 were randomly allocated to the role of Public Official A and the other 11 participants were assigned the role of Public Official B. In each round, each A was randomly matched (with equal probability) with a new B (random strangers). Public Official A was not informed about

---

[3] As discussed by Abbink and Serra (2012), the use of NGOs or charities as recipients of non-embezzled funds is a useful way to model the negative impact of corruption on public well-being. There are, however, two potential issues. First, there may be some loss of control regarding a subject attitude towards a particular NGO or charity. To mitigate this effect, we simply informed the subject that the local NGO would be drawn randomly from a list. Second, subjects' donation behaviour outside the experiment is unknown, if the subject had already given to a charity recently or if he/she decides to be corrupt in the experiment and give later. This may, however, be a non-optimal choice given the multiplicative factor in our experiment.

[4] We chose two different recipients, one for Public Official A and another for Public Official B, in order to make sure that A's donation does not substitute for the donation made by B or influence its marginal effect (see Francois 2000, 2003 for further theorizing as to why such issues may well matter, though in a slightly different context).

the choice made by Public Official B for the first 20 rounds but for the final 20 rounds he or she observed how much Public Official B transferred to Recipient 2. Once all 40 rounds were complete, the subjects were asked to answer a survey that included questions on demographics, socio-economic status, attitudes to and experiences of corruption. Finally, one of the rounds was randomly drawn, and the payments to the participant and the recipient organization were carried out according to the outcome of that round.

3.2 Treatments

Our interest in this paper is in studying if and how the corrupt actions of policy makers (Public Official A) influence others' corrupt behaviour as well as the effectiveness of the anti-corruption mechanism that they choose. Specifically, we wish to see if the level of detection and punishment is less effective in terms of deterring embezzlement when chosen by a corrupt policy maker as opposed to an honest one. To this end, we implemented four experimental treatments, of which three featured a detection and punishment mechanism. Detection automatically implies punishment in that if a public official is detected embezzling from the social fund under his or her control, then that public official forfeits his or her salary in addition to the funds that were embezzled. The first treatment (NoD for No Detection) lacks any scope for detection and punishment and the public officials make their decisions as described above.

The second treatment, which we call Endogenous and Discretionary (ED), gives Public Official A the responsibility of choosing at no cost a probability of detection and punishment which must be selected from the values 0 per cent, 5 per cent, 10 per cent, 15 per cent, 20 per cent, 25 per cent, or 30 per cent. Detection is discretionary in that this mechanism only applies to Public Official B. This setting captures a weak institutional environment in which Public Official A faces no risk when engaging in corruption while Public Official B does. In other words, the principle of equality before the law does not hold. Public Official B acts only after he or she has observed the choices made by Public Official A with respect to the level of detection probability and embezzlement. Detection means that Public Official B loses all earnings for that round.

Treatment three, Endogenous and Non-Discretionary (END), also gives Public Official A the power to select the likelihood of detection (at no cost and from among the same values as above) but this probability applies to both public officials. That is to say that the enforcement of the law is non-discretionary. This is a stronger institutional setting in the sense that equality before the law is a feature but note that the framework is still manipulable. A public official who is detected embezzling loses his/her salary for the round and the amount embezzled in that round. Independent and separate draws are carried out for each public official meaning that in situations where both are corrupt one can be detected and punished while the other is not. Once again Public Official B observes the choices of Public Official A before making his own decision.

The final treatment, Exogenous and Non-Discretionary (XND), exogenously sets the probability of detection at 30 per cent and applies it to both public officials. As with END, independent and separate draws are carried out for each public official. This represents a strong, non-manipulable institutional environment with equality in legal procedure.

Detection and punishment mechanism

The monitoring mechanism functions in a clear and straightforward manner. Once the public officials have made their decisions, the computer generates a random number between 1 and 100. In treatments where both public officials are subject to the mechanism, separate and independent draws are made for each public official. Say a public official opts to keep a positive amount of the social fund for himself and the probability of detection that has been chosen (or has been exogenously imposed) is 30 per cent. If the randomly generated number for that public official falls between 1 and 30 (inclusive) then the public official's decision to embezzle is detected and punished. For that specific round, the public official loses both his salary and the embezzled funds but this does not affect the payoffs in any other round. If the randomly generated number falls between 31 and 100 (inclusive) then the public official in question gets to keep both his salary and the amount kept. The detection and punishment mechanism operates identically in all treatments. The probability value is chosen by Public Official A in the ED and END treatments and is exogenously set at 0 per cent and 30 per cent in the NoD and XND treatments respectively. Table 1 summarizes this procedure for each potential probability value.

Table 1: Details of the detection mechanism

	Numbers are generated between 1 and 100	
Probability values	Randomly generated numbers for which a public official loses both his or her salary and the amount of the social fund kept.	Randomly generated numbers for which a public official retains both his or her salary and the amount of the social fund kept.
0%	Never	Always
5%	1,…,5	6,…, 100
10%	1,…,10	11,…,100
15%	1,…,15	16,…,100
20%	1,…,20	21,…,100
25%	1,…,25	26,…,100
30%	1,…,30	31,…,100

Source: Authors' illustration.

3.3 Participants and payoffs

Across all treatments, 262 subjects participated at the Busara Center for Behavioral Economics, in Nairobi, Kenya. Half took the role of Public Official A and the other half took the role of Public Official B. Sixty-four subjects served in the NoD treatment and 64, 68, and 66 in the ED, END, and XND treatments respectively. Table 2 presents summary statistics for our sample of Public Officials B. They are roughly 22 years old on average and most of them are male. Economics majors make up large proportions of the sample in all of our treatments. The average monthly expenses are around 9500Ksh which is equivalent to around €80. Thus the average earnings from the experiment as described below represent a significant sum to our participants. There are some differences across treatments in some of these characteristics. Though the subjects in ED and END, the treatments that are of particular interest for this paper, are rather similar, we checked that our results are robust to the inclusion of these factors in our regression analysis.

Table 2 also demonstrates that most of our subjects state that they have been asked for a bribe at some point in their lives. In addition to asking about their experiences of corruption, our survey also probes our subjects' attitudes to and understanding of corruption. In terms of perceptions of corruption, 5 per cent of Public Officials B think that a few government officials are involved in corruption, 78 per cent think that some of them are, and the remainder think that all of them are. Most of our subjects (63 per cent) most often hear about corruption in the context of scandals involving politicians and bureaucrats. Twenty-seven per cent of our subjects most often hear about corruption in the context of harassment bribes levelled at ordinary people by government officials and 8 per cent in the context of scandals involving companies and rich individuals. Eighty-nine per cent agree that Kenyan law is such that both bribe takers and givers are acting illegally. One hundred per cent of Public Officials B profess to agree with the statement 'it is always wrong for a government official to take a bribe'. While survey data on an individual's relationship to corruption may be prone to certain biases, these responses suggest that our subjects have an understanding of the practical and moral facets of corruption.

Table 2: Summary statistics—Public Official B characteristics

	NoD	ED	END	XND
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Age	21.69	21.06	21.76	21.45
	(2.07)	(2.44)	(2.15)	(2.28)
Gender (1 if male)	0.53	0.63	0.65	0.85
	(0.50)	(0.50)	(0.49)	(0.36)
Monthly expenses	7,140.63	11,634.38	9,871.18	9,539.42
	(5,127.84)	(17,778.31)	(11,386.99)	(7,712.04)
Economics major	0.44	0.56	0.47	0.76
	(0.50)	(0.50)	(0.51)	(0.44)
Has been asked for a bribe (0 if never)	0.69	0.63	0.74	0.88
	(0.46)	(0.49)	(0.45)	(0.33)
Owns means of transportation	0.03	0.19	0.12	0.12
	(0.17)	(0.40)	(0.33)	(0.33)
Observations	**32**	**32**	**34**	**33**

Source: Authors' calculations.

One period from the 40 was chosen at random to calculate the payoffs. With an exchange rate of 8ECU = 1KSh, the average total earnings (i.e. salary plus embezzlement) for those in the role of Public Official A was 196KSh in the NoD treatment, 208KSh in the ED treatment, 194KSh in the END treatment, and 145KSh in the XND treatment. Public Officials B earned 292KSh in the NoD treatment, 307KSh in the ED treatment, 306KSh in the END treatment, and 185KSh in the XND treatment. In addition, each subject received a fixed payment of 400Ksh for their participation.

The NGOs Green Belt Movement and Impacting Youth Trust (Mathare) served as Recipient 1 and Recipient 2 respectively after being randomly drawn from a list of local NGOs. 48,285KSh were transferred to Recipient 1 and 58,900KSh to Recipient 2 after the experiment had ended. These amounts were calculated by taking the total amount sent to Recipient 1 (Recipient 2) by those in the role of Public Official A (Public Official B) using one randomly determined period per subject and an exchange rate of 8ECU = 1KSh.[5]

4 Main results

In this paper, we are interested in explaining the corrupt behaviour of Public Official B, while Boly and Gillanders (2016) analyse Public Official A's behaviour. We also restrict our analysis to the first 20 rounds, given that in the final 20 rounds, Public Official A was informed about Public Official B's embezzlement decision. We analyse the data using summary statistics and statistical tests (mainly Mann-Whitney tests with individual average choices as independent units of observations), followed by regression analyses. Two main variables are of interest: the likelihood that a Public Official B is corrupt and the amount embezzled by corrupt Public Officials B.

Our reading of the literature leads us to expect three distinct effects. Firstly, we expect to see a deterrence effect. Such an effect will be evident if we find a downward-sloping relationship

[5] Each participant was notified once the funds were transferred. The participants also received a text message notifying them that there were receipts and formal letters of NGO payments available for viewing and collection at Busara's offices if they so wished.

between the level of detection probability and corrupt behaviour. Secondly, we expect to see a peer effect whereby a Public Official B who witnesses corrupt behaviour on the part of a Public Official A will follow suit. A difference in the average level of corrupt outcomes by the type of Public Official A will be evidence of this peer effect. Finally, we are interested in the possibility that policies originating from a corrupt source are less effective than those promulgated by an honest policy maker. Such a legitimacy effect could present itself in two ways. Firstly, if the slope of the deterrence effect differs by type of Public Official A we would have evidence of a legitimacy effect. The effect of changes in detection would be different depending on the behaviour of Public Official A. Secondly, and conceptually equivalently, differences in the overall marginal effects of Public Official A's type at each level of detection probability derived from models with interactions between detection and type would be evidence of the same policy having different effects depending on the behaviour of the policy maker.

4.1 Share of corrupt decisions by Officials B

Summary statistics on the average share of corrupt decisions made by Public Official B are given in Table 3. Note that individual average choices are used as independent units of observations. The average shares of corrupt decisions in the NoD, ED, END, and XND treatments are respectively 78 per cent, 88 per cent, 91 per cent, and 79 per cent. Relative to the NoD treatment, corruption is not significantly higher in the ED treatment (p-value = 0.61, two-sided Mann-Whitney) or in the END treatment (p-value = 0.14, two-sided Mann-Whitney). Compared to the XND treatment, we find that corruption is significantly greater in the END treatment (p-value = 0.0420, two-sided Mann-Whitney) but not in the ED treatment. No significant difference is found between the NoD and the XND treatments (p-value = 0.66, Mann-Whitney) or between the ED and the END treatment (p-value = 0.31, Mann-Whitney).

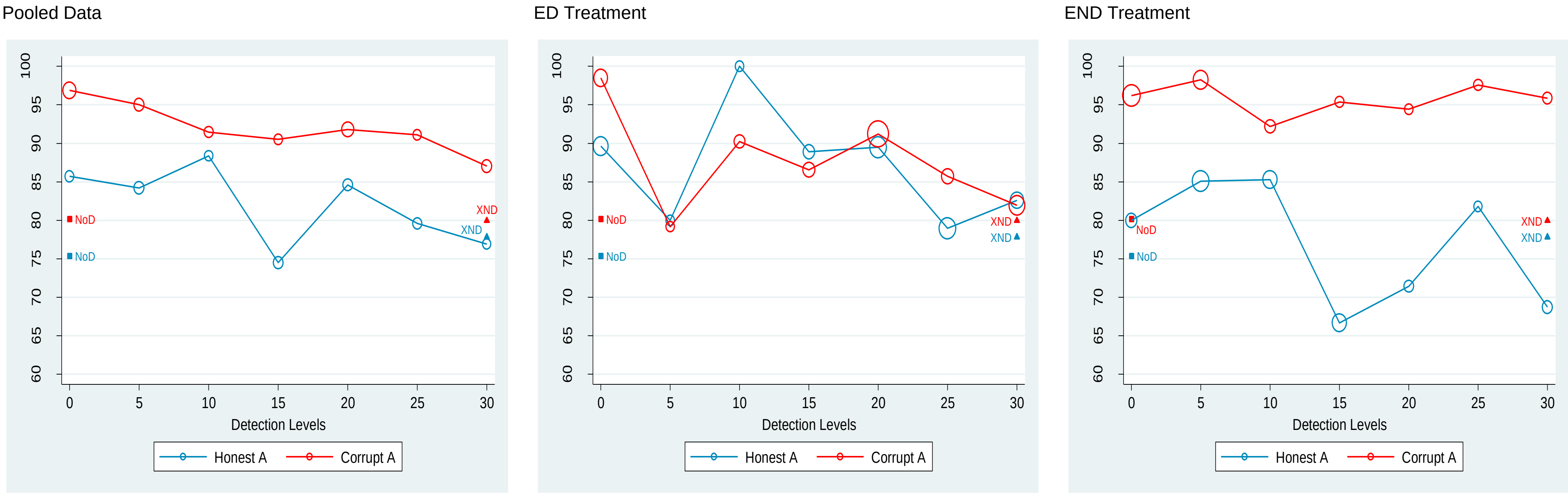
Table 3: Average choices of Public Official B by treatment

	(1)	(2)	(3)	(4)
	NoD	ED	END	XND
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
B's behaviour (corrupt=1)	0.78	0.88	0.91	0.79
	(0.35)	(0.19)	(0.16)	(0.30)
Amount kept by B	1,203.56	1,487.65	1,493.75	1,198.49
	(772.48)	(512.89)	(590.52)	(689.88)
Subjects	**32**	**32**	**34**	**33**

Source: Authors' calculations.

Figure 1 plots, overall and for each treatment, the percentage of Public Officials B who made corrupt decisions at each point on our detection scale. It also includes reference points for the treatments where this probability was exogenously set (NoD and XND treatments). The share of corrupt decisions conditional on Public Official A being corrupt is indicated in red while the share of corrupt decisions conditional on A being honest is indicated in blue. The pooled data suggest that deterrence is effective in that no matter the type of Public Official A we see a downward slope indicating that higher levels of detection lower the likelihood of a corrupt Public Official B. However, the type of Public Official A does seem to matter in that the line for honest Public Official As is always below the line for corrupt Public Officials A.

Figure 1: Share of corrupt Public Official B's—by detection level and Public Official A's type



Note: The bubble size corresponds to the percentage of time a given level of detection was endogenously chosen.

Source: Authors' calculations.

In the ED treatment, a negative relationship between the probability of detection and the share of corrupt Public Officials B can be observed and the relationship appears similar for each type of Public Official A. This suggests that the decision by Public Official A to be corrupt or honest does not influence the effectiveness of the chosen level of detection in terms of deterring Public Official B from embezzlement. In contrast, in our END treatment, there is a clear difference depending on whether Public Official A is corrupt or honest. At each point on our detection scale, the share of corrupt Public Officials B is appreciably lower when the probability has been chosen by an honest Public Official A as opposed to a corrupt Public Official A. In addition, increasing the probability does not seem to dissuade Public Official B from being corrupt when that probability has been chosen by a corrupt Public Official A. When Public Official A is honest, we do see some evidence of a downward slope. These differences are indicative of peer and legitimacy effects.

We now proceed to a regression analysis to test the statistical significance and magnitude of these apparent effects. Controls such as age, gender, monthly expenses (in log), and economics as major are typically found to be insignificant and are therefore not included in the regressions.[6] We use a random-effects Logit model to analyse Public Official B's decision to embezzle or not. The results are in line with our graphical analysis for the most part.

Pooled data

In our regression analyses, we estimate equations of the general form:

$$y_{it} = \alpha + \beta END + \rho Honest\_A_{it} + \delta Detection_{it} + \mu_i + \varepsilon_{it} \quad (1)$$

$$y_{it} = \alpha + \beta END + \gamma_1(Honest\_A_{it} \times Detection_{it}) + \gamma_2(Corrupt\_A_{it} \times Detection_{it}) + \mu_i + \varepsilon_{it} \quad (2)$$

$$y_{it} = \alpha + \beta END + \rho Honest\_A_{it} + \delta Detection_{it} + \gamma(Honest\_A_{it} \times Detection_{it}) + \mu_i + \varepsilon_{it} \quad (3)$$

$$y_{it} = \alpha + \beta END + \rho Honest\_A_{it} + \delta Dummy\_Detection_{it} + \mu_i + \varepsilon_{it} \quad (4)$$

$$y_{it} = \alpha + \beta END + \gamma Dummy\_(Honest\_A_{it} \times Detection_{it}) + \mu_i + \varepsilon_{it} \quad (5)$$

y_{it} is the dependent variable. It is either a dummy variable which equals 1 when Official B embezzles (0 otherwise), or the amount, between 0 and 2,280ECU, embezzled by Official B (see Section 4.2). The variable END is a dummy variable which equals 1 for the END treatment and 0 otherwise. $Honest\_A_{it}$ is a dummy variable which equals 1 when Official A is honest and 0 otherwise. $Detection$ is the level of detection chosen by Official A. It is treated as a continuous variable in Equations 1–3 while a dummy is created for each level of detection in Equations 4–5. $Corrupt\_A_{it}$ is a dummy variable which equals 1 when Official A is corrupt and 0 otherwise.

[6] The results are quantitatively and qualitatively similar if these controls are included.

$Honest\_A_{it} \times Detection_{it}$ and $Corrupt\_A_{it} \times Detection_{it}$ are the interactions of Public Official A's behaviour with the level of detection that he/she chooses. The constant term, individual effect term, and the error term are respectively α, μ_i and ε_{it}. The subscript i is for individual subjects and t denotes the round.

In Table 4, we pool the data from the ED and END treatments and start by looking at the main effects of Public Official A's behaviour and the level of detection probability (which is treated as continuous in columns 1 to 3) on the likelihood that Public Official B acts corruptly. In Table 4, column 1, we include only the main effects in the regression (Equation 1**Error! Reference source not found.**) and find a significant and negative relationship between honest behaviour by Public Official A and the likelihood of embezzlement by Public Official B (at the 1 per cent level). We also find that deterrence has a negative and statistically significant effect on the likelihood that Public Official B acts corruptly. This deterrence effect, which is in line with much of the experimental literature outlined above, is evident in Figure 1, panel A as the downward slope in both the red and blue lines. Column 4, in which we employ dummies for each detection level (Equation 4), supports this conclusion as several levels of deterrence have negative effects on Public Official B's likelihood to embezzle. In particular the 20 per cent and 30 per cent levels are statistically significant at the 5 per cent and 1 per cent levels respectively. Our pooled data thus suggests is the presence of a simple peer effect and of a deterrence effect.

Table 4: Likelihood of embezzlement by Public Official B—pooled data (Logit)

	Detection Levels—continuous			Detection levels—dummy	
	(1)	(2)	(3)	(4)	(5)
VARIABLES	Main effects	Interaction only	Full model	Main effects	Interaction only
END	0.580	0.597	0.586	0.599	0.615
	[0.686]	[0.705]	[0.694]	[0.698]	[0.713]
A’s behaviour (honest=1)	-1.144***		-0.794*	-1.191***	
	[0.244]		[0.444]	[0.249]	
Detection level	-0.030**		-0.022		
	[0.012]		[0.015]		
A’s behaviour=0 # detection level		-0.008			
		[0.012]			
A’s behaviour=1 # detection level		-0.067***	-0.023		
		[0.015]	[0.024]		
Detection level=5				-0.512	
				[0.470]	
Detection level=10				-0.698	
				[0.477]	
Detection level=15				-0.765*	
				[0.451]	
Detection level=20				-0.983**	
				[0.443]	
Detection level=25				-0.438	
				[0.474]	
Detection level=30				-1.273***	
				[0.435]	
A’s behaviour=0 # detection level=5					-0.393
					[0.622]
A’s behaviour=0 # detection level=10					-0.690
					[0.603]
A’s behaviour=0 # detection level=15					-0.461
					[0.614]
A’s behaviour=0 # detection level=20					-0.815
					[0.549]
A’s behaviour=0 # detection level=25					-0.230
					[0.631]
A’s behaviour=0 # detection level=30					-0.900*
					[0.542]
A’s behaviour=1 # detection level=0					-0.763
					[0.667]
A’s behaviour=1 # detection level=5					-1.471**
					[0.636]
A’s behaviour=1 # detection level=10					-1.379*
					[0.735]
A’s behaviour=1 # detection level=15					-1.867***
					[0.635]
A’s behaviour=1 # detection level=20					-1.950***
					[0.681]
A’s behaviour=1 # detection level=25					-1.430**
					[0.677]
A’s behaviour=1 # detection level=30					-2.812***
					[0.701]
Constant	4.255***	3.883***	4.145***	4.529***	4.392***
	[0.578]	[0.576]	[0.592]	[0.635]	[0.688]
Lnsig2u constant	1.641***	1.710***	1.670***	1.670***	1.721***
	[0.335]	[0.334]	[0.336]	[0.335]	[0.337]
Observations	**1,320**	**1,320**	**1,320**	**1,320**	**1,320**
Subjects	**66**	**66**	**66**	**66**	**66**

Note: Standard errors in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: Authors’ calculations.

The regression framework allows us to go deeper and study the interaction of Public Official A's behaviour and the level of detection that he chooses. In column 2 of Table 4 (corresponding to Equation 2), we include only interaction terms between Public Official A's behaviour and the level of detection.[7] The level of detection has no effect on the likelihood that Public Official B acts corruptly when chosen by a corrupt Public Official A. In contrast, when chosen by an honest Public Official A, a greater chance of detection significantly decreases the likelihood of embezzlement by Public Official B (at the 1 per cent level). The difference between the coefficients for each type of Public Official A is significant at the 1 per cent level (p-value = 0.000, chi2). This is consistent with a legitimacy effect.

Column 3 of Table 4 includes the main effects and the interaction effect (Equation 3). The reference group is a Public Official B in the ED treatment who is paired with a corrupt Public Official A who has selected a zero probability of detection. The coefficient for 'detection level' is negative suggesting that detection is effective when chosen by an honest Public Official A but the effect is not significant at traditional levels. The coefficient for the dummy representing an honest Public Official A is negative (and significant at the 10 per cent level) indicating that the likelihood of embezzlement by Public Official B is higher when facing a corrupt Public Official A instead of an honest Public Official A. Graphically, this would mean that the line of predicted Logit values for corrupt Public Officials A will lie above the line for their honest counterparts.

The insignificant interaction term might seem to suggest that the marginal effect of Public Official A's behaviour does not depend on the level of detection. However, the marginal effects of Public Official A's type may be significantly different for different detection levels. Indeed, note that the coefficient for 'A's behaviour' in column 3 of Table 4 gives the overall effect of an Honest A when 'detection level' equals 0. Since detection level is a continuous variable, it takes many other values than 0. To understand the overall effects of 'A's behaviour' on embezzlement, we plug in different values of *Detection Level* into Equation 3 (column 3 in Table 4). This allows us to see, for each level of detection, how the likelihood of embezzlement (the dependent variable) changes depending on Public Official A's type (corrupt or honest). The first column of Table 5 presents the results on the differences in the overall average marginal effects between corrupt and honest Officials A. They suggest that the probability of embezzlement by Public Official B is significantly higher when paired with a corrupt Public Official A than with an honest Public Official A, at all levels of detection.[8] That is to say that the same policy yields different results depending on the behaviour of the policy maker. This is consistent with the results presented in column 2 of Table 4. The models in columns 2 and 3 of Table 4 thus also provide evidence that legitimacy matters for the effectiveness of deterrence as anti-corruption policy.

---

[7] In column 2 of Table 4 (Interaction only), we have two continuous variables (CorruptA*Detection and HonestA*Detection) as detection is treated as a continuous variable. The standard interpretation of the constant in a regression equation is the expected mean value of Y (dependent variable) when all other explanatory variables are 0. 'CorruptA*Detection' and 'HonestA*Detection' are simultaneously 0 when detection = 0 for either type of public official A. As a result, the constant does not represent a specific reference group but a mean for corrupt and honest A when detection is 0. The same applies to columns 3 and 4 in Table 6, column 2 in Table 8, and columns 3 and 4 of Table 9.

[8] The difference in marginal effects is: 'Marginal Effect of Honest A' - 'Marginal Effect of Corrupt A'.

Table 5: Differences in average marginal effects between corrupt and honest Officials A—by detection level

	(1)	(2)	(3)	(4)	(5)	(6)
	Share of corrupt B (Logit)			Amount embezzled by B (Tobit)		
	Pooled	ED	END	Pooled	ED	END
Detection level at 0	-0.79*	-0.44	-0.46	-248.77**	-104.26	-267.93*
	[0.44]	[0.74]	[0.61]	[119.23]	[202.79]	[143.09]
Detection level at 5	-0.91***	-0.45	-1.04**	-232.03**	-60.79	-341.52***
	[0.35]	[0.59]	[0.46]	[93.34]	[160.24]	[110.24]
Detection level at 10	-1.02***	-0.45	-1.63***	-215.29***	-17.31	-415.11***
	[0.28]	[0.46]	[0.38]	[74.02]	[124.60]	[90.43]
Detection level at 15	-1.13***	-0.45	-2.22***	-198.55***	26.16	-488.70***
	[0.24]	[0.37]	[0.41]	[67.20]	[103.27]	[92.47]
Detection level at 20	-1.25***	-0.46	-2.81***	-181.81**	69.63	-562.29***
	[0.27]	[0.35]	[0.52]	[76.30]	[105.35]	[115.21]
Detection level at 25	-1.36***	-0.46	-3.40***	-165.08*	113.10	-635.88***
	[0.34]	[0.42]	[0.69]	[96.95]	[129.72]	[149.47]
Detection level at 30	-1.48***	-0.47	-3.99***	-148.34	156.57	-709.47***
	[0.43]	[0.54]	[0.87]	[123.48]	[166.88]	[189.10]
Observations	**1,320**	**640**	**680**	**1,320**	**640**	**680**

Note: Standard errors in brackets. The difference in marginal effects is: ‘Marginal effect of honest A’ - ‘Marginal effect of corrupt A’. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: Authors’ calculations.

In column 5 of Table 4 (corresponding to Equation 5), we employ dummies for each detection level and include only the interaction terms. Thus we take a cell-means model approach in which the baseline cell is a corrupt Public Official A choosing a zero level of detection and all other cells are compared to this baseline. All levels of detection (except level 0) have a statistically significant and negative effect on Public Official B’s likelihood to be corrupt when chosen by an honest Public Official A. In contrast, the coefficients for detection levels chosen by a corrupt Public Official A are all insignificant. A joint equality test of the coefficients for all levels of detection yields significant results (p-value = 0.000, chi2) suggesting that Public Official B is more likely to be corrupt when detection is chosen by a corrupt Public Official A than when chosen by an honest Public Official A. Once again this is evidence of a legitimacy effect.

Thus our pooled data points to a role for deterrence, peer effects, and legitimacy in the fight against corruption. We will discuss the policy implications of these results in the concluding section, but first we examine whether the institutional framework in which our public officials operate alters the importance of these effects.

Individual treatments data

We already noted above that Figure 1 suggests that the effects of interest are heterogeneous across treatments. Table 6 repeats our core regression analysis for the ED and END treatments individually (Equations 1–3). We begin again by looking at the main effects in columns 1 and 2. In the ED treatment, deterrence seems to be at work, as the coefficient on detection level is negative and significant at the 1 per cent level (column 1). However, we do not see evidence of a peer effect in this institutional setting. Things are very different in the institutional setting represented by the END treatment where a peer effect appears to be at work, with the coefficient for Public Official A’s behaviour being significant at the 1 per cent level, while the coefficient for detection level is not significant (column 2).

Table 6: Likelihood of embezzlement by Public Official B—by treatment (Tobit)

	Main effects		Interaction only		Full model	
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	ED	END	ED	END	ED	END
A's behaviour (honest=1)	-0.457	-1.882***			-0.441	-0.455
	[0.347]	[0.369]			[0.742]	[0.610]
Detection level	-0.046***	-0.012			-0.046**	0.043
	[0.017]	[0.019]			[0.020]	[0.028]
A's behaviour=0 # detection level			-0.040**	0.054**		
			[0.017]	[0.024]		
A's behaviour=1 # detection level			-0.060***	-0.089***	-0.001	-0.118***
			[0.021]	[0.024]	[0.036]	[0.042]
Constant	4.280***	5.003***	4.144***	4.540***	4.275***	4.656***
	[0.655]	[0.726]	[0.639]	[0.767]	[0.681]	[0.774]
Lnsig2u constant	1.578***	1.684***	1.581***	1.955***	1.578***	1.901***
	[0.454]	[0.499]	[0.454]	[0.491]	[0.454]	[0.499]
Observations	**640**	**680**	**640**	**680**	**640**	**680**
Subjects	**32**	**34**	**32**	**34**	**32**	**34**

Note: Standard errors in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: Authors' calculations.

Columns 3 and 4 of Table 6 show that the interactive effect of the detection level and Public Official A's behaviour also varies between the ED and the END treatments. In the ED treatment, detection and punishment significantly decreases the likelihood of Public Official B being corrupt when chosen by a corrupt Public Official A. Likewise, there is a significant negative effect when the level of detection is chosen by an honest Public Official A. However, the difference in the marginal effect of the detection level by type of Public Official A is not statistically significant (p-value = 0.238, chi2). In the END treatment, however, detection and punishment appear to increase the likelihood of Public Official B being corrupt when chosen by a corrupt Public Official A and this effect is statistically significant at the 5 per cent level. In contrast, there is a significant negative effect when the probability of detection is chosen by an honest Public Official A. The difference in the effect of the detection level chosen by corrupt or honest Public Official As is significant at the 1 per cent level (p-value = 0.000, chi2). Such a result suggests that, under certain conditions, an anti-corruption policy based on detection and punishment can be counter-productive when chosen by a policy maker who is herself corrupt. This result is also consistent with that of Chen (2013), although in a different context.

The full model presented in column 5 of Table 6 shows that there is no significant peer effect in the ED treatment but that there is a deterrence effect. Table 5, column 2, further confirms that there is no significant effect of Public Official A's type at any detection level. Thus we conclude that in the particular institutional framework captured by the ED treatment, monitoring and punishment are effective and the behaviour of the rule maker is irrelevant. These conclusions are in line with the graphical analysis in panel A of Figure 1 and the findings of the literature on exogenously given detection levels that began with Abbink et al. (2002).

In the END treatment (Table 6, column 6), we find no evidence of a peer effect and we can also see that detection levels chosen by a corrupt Public Official A have no significant effect on the likelihood of embezzlement by Public Official B. However, the negative and significant interaction term suggests that choices made by honest Public Officials A tend to decrease the likelihood of embezzlement by Public Official B. This is consistent with the idea of a legitimacy effect. Once again these conclusions are fully consistent with the graphical analysis presented in Figure 1. They are also confirmed in Table 5, column 3. Public Official B embezzles with a significantly higher probability (at conventional 1 per cent or 5 per cent levels), when facing a corrupt rather than an honest Public Official A, at all non-zero levels of detection.

Our results thus indicate that in the END treatment, a higher probability of detection and punishment can effectively deter corruption, but only when the policy is put in place by a policy maker who is himself 'clean'. Thus, in institutional settings in which equality before the law is observed, it is vital that those at the top set the right 'tone', if endogenously chosen anti-corruption measures of this type are to be effective. This raises the question as to whether exogenously imposed policies will fare any better in the face of corrupt leaders.

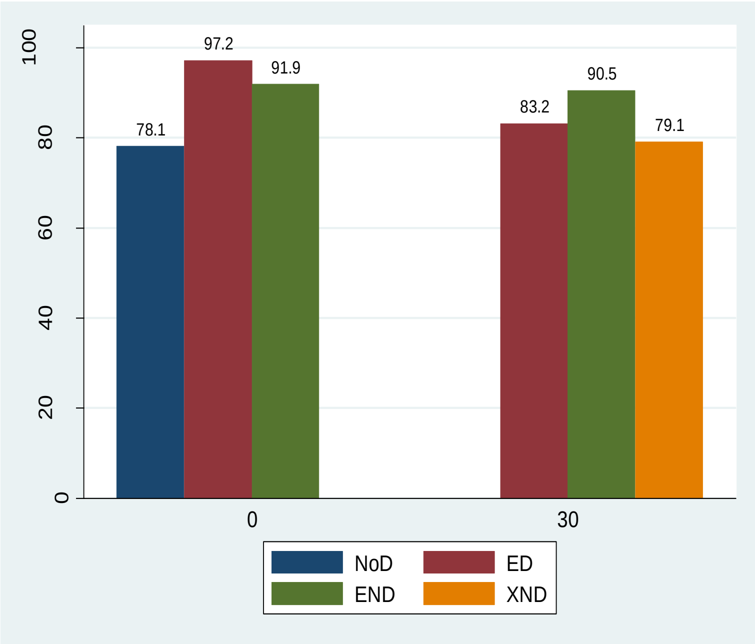
Externally vs internally imposed detection

Our design allows us to examine this further important issue in the fight against corruption. Namely, do rules that are externally imposed have the same effectiveness as ones that are chosen by the members of a 'society'? Furthermore, does the answer to this question depend on the legal structure in place and the corrupt actions of local policy makers? We analyse these questions by focusing on specific detection levels (namely 0 and 30) across all four treatments.

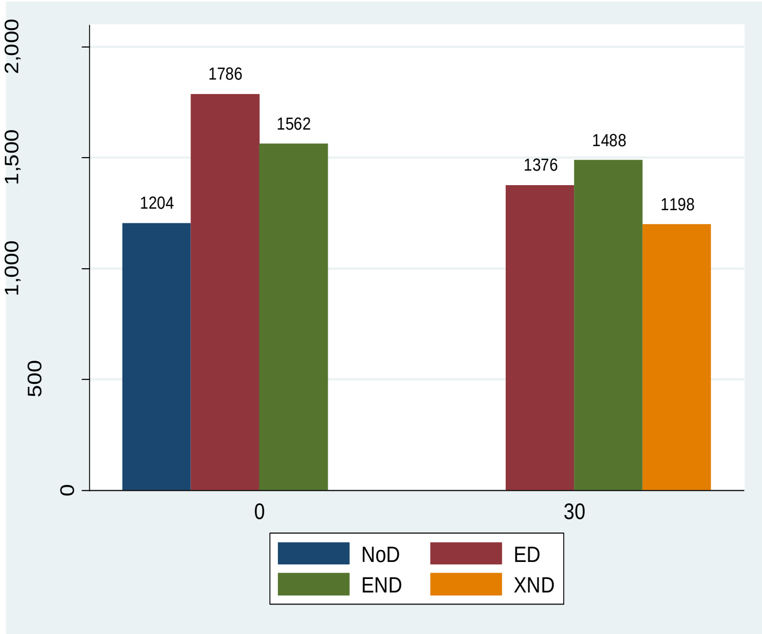
Figure 2.A shows that when the detection probability is 0, the share of corrupt decisions by Public Officials B in the ED and END treatments are 97.2 per cent and 91.9 per cent respectively, compared to 78.1 per cent in the NoD treatment. The regression results in Table 7, panel A, column 1 indicate that holding the detection level at 0, the likelihood of embezzlement is significantly higher in the ED and END treatments compared to the NoD treatment. We also find evidence of peer effects when there is no detection in that honesty by Public Official A significantly decreases the likelihood of embezzlement by Public Official B. These results are in line with the idea that an explicit choice of 0 probability of detection by Public Official A (in the ED and END treatments) can be interpreted by Public Official B more as tolerance for corruption than as a sign of trust in his honesty. In Table 7, panel A, column 2, we introduce interactions between Public Official A's behaviour and treatments. There is no significant difference generated by the type of Public Official A in the NoD and the END treatment though the difference is significant at the 10 per cent level in the ED treatment.

Figure 2: Externally vs internally imposed detection

A. Average share of corrupt decisions by Official B



B. Average amount embezzled by Official B



Source: Authors' calculations.

Perhaps more interestingly from the perspective of external development and stability agencies is the fact that when the probability of detection is 30 per cent, the shares of corrupt decisions are 83.2 per cent, 90.5 per cent, and 79.1 per cent in the ED, END, and XND treatments respectively. However, we find no statistically significant differences between the ED and the XND treatments and the END and the XND treatments when the detection probability is 30 per cent (Table 7, panel A, column 3). The results do show that a strong endogenously chosen anti-corruption policy leads to better outcomes in terms of corruption relative to an equally strong exogenously given policy, only when we have a situation with equality before the law and an honest policy maker (Table 7, panel A, column 4). This lends some support to the idea that external stakeholders can play a role in curbing corruption when they face entrenched corruption at the domestic policy-making level.

Table 7: Exogenous vs endogenous detection—levels 0 and 30%

Panel A: Likelihood of embezzlement by Public Official B (corrupt=1)

	Detection level 0		Detection level 30	
	(1)	(2)	(3)	(4)
	Main effects	Full model	Main effects	Full model
	Baseline NoD		Baseline XND	
ED	3.254**	6.024**	0.477	0.779
	[1.596]	[2.596]	[0.890]	[0.979]
END	2.426**	2.657**	1.345	3.424**
	[1.166]	[1.242]	[1.026]	[1.550]
A's behaviour (honest=1)	-0.696**	-0.503	-0.204	0.037
	[0.334]	[0.355]	[0.266]	[0.286]
ED # A's behaviour=1		-4.455*		-0.817
		[2.468]		[1.041]
END # A's behaviour=1		-0.481		-4.473***
		[1.377]		[1.714]
Constant	3.504***	3.569***	2.726***	2.693***
	[0.750]	[0.806]	[0.569]	[0.595]
Lnsig2u constant	2.684***	2.817***	2.094***	2.210***
	[0.370]	[0.396]	[0.340]	[0.353]
Observations	913	913	830	830
Subjects	89	89	93	93

Panel B: Amount embezzled by Public Official B

	Detection level 0		Detection level 30	
	(1)	(2)	(3)	(4)
	Main effects	Full model	Main effects	Full model
	Baseline NoD		Baseline XND	
ED	1,000.609***	1,043.837***	192.465	187.038
	[346.920]	[354.022]	[303.787]	[311.795]
END	668.731**	666.452**	460.659	619.453*
	[303.847]	[309.937]	[321.104]	[336.413]
A's behaviour (honest=1)	-154.055**	-142.786**	-131.669*	-105.223
	[63.329]	[68.027]	[78.939]	[84.612]
ED # A's behaviour=1		-150.132		42.090
		[228.438]		[293.695]
END # A's behaviour=1		28.880		-573.892
		[294.283]		[358.110]
Constant	1,210.171***	1,205.314***	1,242.393***	1,231.762***
	[214.399]	[215.139]	[201.432]	[201.371]
Sigma_u constant	1,173.164***	1,175.856***	1,111.448***	1,108.905***
	[112.648]	[113.163]	[111.512]	[111.095]
Sigma_e constant	688.940***	688.900***	921.762***	921.406***
	[23.065]	[23.066]	[32.947]	[32.937]
Observations	913	913	830	830
Subjects	89	89	93	93

Note: Standard errors in brackets: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

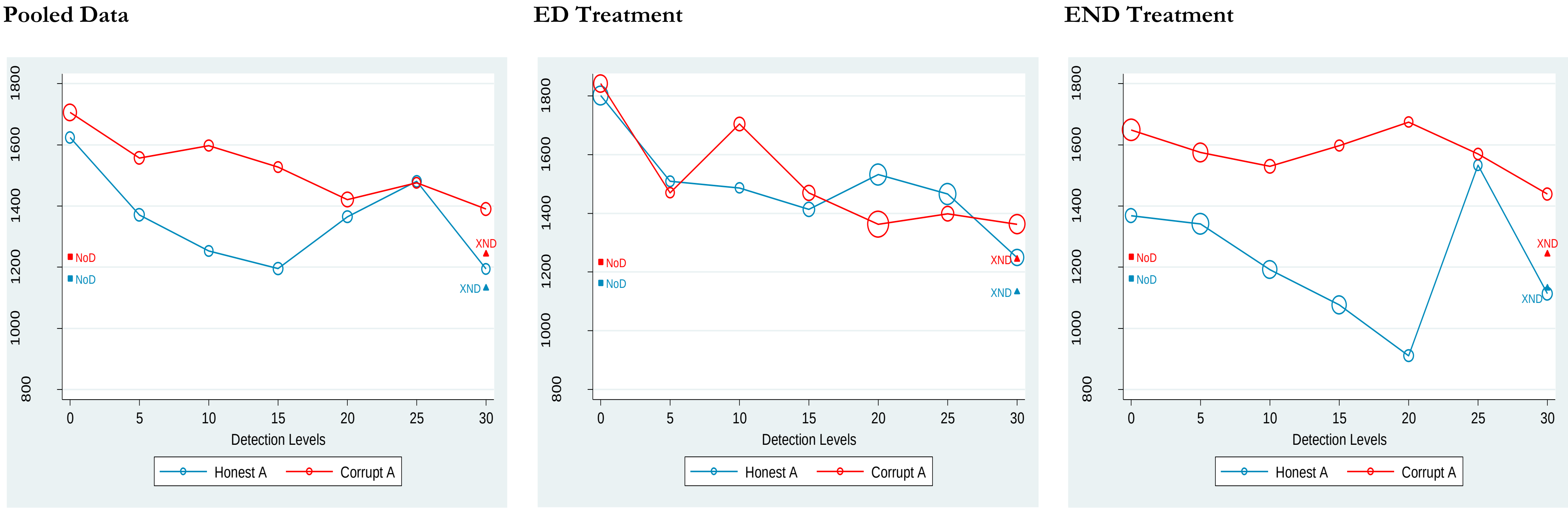
Source: Authors' calculations.

4.2 Amount embezzled by Public Official B

Next we analyse the amount embezzled by Public Official B. Especially in contexts where some level of corruption is expected, the factors outlined above could operate on the decision regarding the extent of corruption as well as the decision to be corrupt or not. That is to say that the tone at the top could lead to people embezzling more or less even though the probability of detection and punishment is the same no matter the level of their embezzlement. The analysis here mirrors the one conducted in Section 4.1. The amounts kept by Public Official B in the NoD, ED, END, and XND treatments are respectively 1,204ECU, 1,488ECU, 1,494ECU, and 1,199ECU (see Table 3). Relative to the NoD treatment, the amounts embezzled, in the ED and END treatments, are both not significantly higher at conventional levels (p-values of 0.17 and 0.19, respectively, two-sided Mann-Whitney). In contrast, compared to XND, we find that the amounts embezzled increase significantly both in the ED and END treatments at the 10 per cent significance level (p-value = 0.09 for both, two-sided Mann-Whitney). No significant difference is found between the NoD and the XND treatments (p-value = 0.93, Mann-Whitney); or between the ED and END treatments (p-value = 0.76, Mann-Whitney).

We now proceed to disaggregate the amount embezzled by Officials B according to Public Official A's behaviour (honest or corrupt) and the detection level chosen by Public Official A. Figure 3 carries out the same exercise as Figure 1, for the extent of Public Official B's embezzlement. Once again, it should be noted that since, for a given probability of detection, embezzling 1ECU is as likely to result in punishment as embezzling 2,280ECU, it is not obvious that higher probabilities should lead to an individual embezzling a lower amount. However, we observe a negative relationship between the level of detection and the amount embezzled in the pooled data and in the ED treatment. Such evidence of deterrence is observed whether Public Official A is corrupt or honest. The pooled data picture is consistent with a peer effect and a legitimacy effect, though once again these effects appear to be unimportant in the ED treatment. The END sample shows clear evidence of a peer effect and a legitimacy effect in that each endogenously chosen level of detection probability (with the possible exception of the 25 per cent level) gives rise to a lower average amount embezzled when chosen by an honest as opposed to a corrupt Public Official A. For both types, we see some evidence of a deterrent effect of monitoring and punishment.

Figure 3: Mean embezzlement of Public Officials B—by detection level and Public Officials A type



Note: The bubble size corresponds to the percentage of time a given level of detection was endogenously chosen.

Source: Authors' calculations.

Next we verify that the broad patterns observed in Figure 3 are statistically significant using regression analysis. As the degree of embezzlement had to be chosen from a restricted range of values (between 0 and 2,280ECU), we employ a random-effects two-sided Tobit model for our analysis of embezzlement by Public Official B.[9] The results confirm our graphical analysis for the most part.

Pooled data

In Table 8, we pool the data from the ED and END treatments and start in column 1 by looking at the main effects of Public Official A's behaviour and of the probability of detection on the amount embezzled (Equation 1). Our regression shows a statistically significant and negative relationship between the detection level and the amount embezzled by Public Official B (at the 1 per cent level). The magnitude of this deterrence effect is a roughly 11ECU decrease in embezzlement per additional 1 per cent probability of detection. There is also evidence of a peer effect as the coefficient on the dummy variable capturing an honest Public Official A is significant and negative. This magnitude of this effect is also meaningful. When facing a corrupt Public Official A, Public Official B embezzles an additional 200ECU on average. Column 4 of Table 8 employs a dummy for each detection level and concludes that there is a meaningful deterrence effect at all levels of detection. This result supports the conclusion that detection is effective in curbing the amount embezzled.

[9] Again, we exclude controls as they are not significant and the results remain similar if included.

Table 8: Amount embezzled by Public Official B—pooled data (Tobit)

	Detection levels—continuous			Detection levels—dummy	
	(1)	(2)	(3)	(4)	(5)
VARIABLES	Main effects	Interaction only	Full model	Main effects	Interaction only
END	9.813	13.163	8.726	22.636	44.585
	[190.606]	[191.213]	[190.567]	[192.130]	[192.371]
A's behaviour (honest=1)	-199.841***		-248.772**	-191.624***	
	[67.160]		[119.227]	[67.339]	
Detection level	-10.807***		-11.613***		
	[3.035]		[3.442]		
A's behaviour=0 # detection level		-8.718***			
		[3.150]			
A's behaviour=1 # detection level		-17.004***	3.348		
		[4.245]	[6.738]		
Detection level=5				-304.588***	
				[100.041]	
Detection level=10				-363.048***	
				[109.614]	
Detection level=15				-397.836***	
				[110.310]	
Detection level=20				-404.543***	
				[100.895]	
Detection level=25				-255.449**	
				[114.025]	
Detection level=30				-428.595***	
				[103.949]	
A's behaviour=0 # detection level=5					-257.555**
					[113.599]
A's behaviour=0 # detection level=10					-263.998**
					[126.544]
A's behaviour=0 # detection level=15					-247.005*
					[131.626]
A's behaviour=0 # detection level=20					-383.438***
					[111.624]
A's behaviour=0 # detection level=25					-275.637**
					[132.903]
A's behaviour=0 # detection level=30					-386.808***
					[118.230]
A's behaviour=1 # detection level=0					1.250
					[172.500]
A's behaviour=1 # detection level=5					-481.615***
					[156.820]
A's behaviour=1 # detection level=10					-657.058***
					[177.086]
A's behaviour=1 # detection level=15					-760.104***
					[166.469]
A's behaviour=1 # detection level=20					-464.827***
					[167.930]
A's behaviour=1 # detection level=25					-271.534
					[176.402]
A's behaviour=1 # detection level=30					-575.989***
					[180.288]
Constant	1842.657***	1791.507***	1854.575***	1973.808***	1922.339***
	[147.004]	[146.195]	[148.937]	[153.315]	[157.141]
Sigma_u constant	734.325***	736.813***	734.115***	737.499***	737.578***
	[71.838]	[72.048]	[71.822]	[72.029]	[72.015]
Sigma_e constant	947.854***	949.125***	947.769***	942.282***	939.372***
	[26.344]	[26.384]	[26.341]	[26.175]	[26.090]
Observations	**1,320**	**1,320**	**1,320**	**1,320**	**1,320**
Subjects	**66**	**66**	**66**	**66**	**66**

Note: Standard errors in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: Authors' calculations.

In Table 8, column 2, we include only the interaction terms between Public Official A's behaviour and detection (Equation 2). We find that the decision made by Public Official A (to be corrupt or honest) has an effect on the amount embezzled by Public Official B. While detection decreases the amount embezzled regardless of whether it was chosen by a corrupt or an honest Public Official A, we find that the effects of detection levels chosen by an honest Public Official A (coefficient of -17.0) are almost double those of detection probabilities chosen by a corrupt Public Official A (coefficient of -8.7). This difference is significant at the 5 per cent level (p-value = 0.0293, chi2) thereby suggesting a legitimacy effect.

The results of the full model (including the main effects and interaction term) are shown in column 3 of Table 8. Here the baseline group is a Public Official B in the ED treatment encountering a corrupt Public Official A who has chosen detection level 0. The coefficient on the level of detection is negatively signed and statistically significant at the 1 per cent level. The coefficient on Public Official A's behaviour is negative and significant at the 5 per cent level, suggesting that on average Public Official B embezzles a lower amount when facing an honest Public Official A. To see whether the effects for honest and corrupt Public Officials A are significantly different at different detection levels, we compute the difference in the average marginal effect of Public Official A's type by detection level. The results, which are presented in Table 5, column 4, suggest that facing a corrupt Public Official A rather than an honest one significantly increases the amount embezzled by Public Official B at all levels of detection except the highest one (i.e. 30 per cent).

Column 5 of Table 8 uses dummies for each detection level (Equation 4) and finds that all levels of detection have a significant negative effect on the amount embezzled by Public Official B when Public Official A is honest. Most levels are also significant whether Public Official A acts corruptly or honestly, using a corrupt official choosing 0 detection level as a baseline (Equation 5). Considering the coefficient for corrupt and honest Officials A, a joint equality test of the coefficients for each level of detection rejects the null hypothesis of joint equality (p-value = 0.0314, chi2). In essence, despite the presence of deterrence for both types, Public Official B is likely to embezzle more funds when Public Official A (the policy maker) is corrupt than when he/she is honest. This result implies that legitimacy can play a significant role in the fight against embezzlement.

Individual treatments data

Table 9 shows that the effects observed for the average amount of embezzlement vary depending on the treatment. That is to say that the institutional framework, specifically the presence of procedural equality before the law, determines the importance of peer, deterrence, and legitimacy effects. We begin by looking at the main effects (Equation 1). The first column of Table 9 tells us that in the ED treatment, where equality before the law is not observed, there is a deterrence effect as the coefficient for detection level is negative and significant at the 1 per cent level. However, the coefficient on Public Official A's behaviour is not significant, leaving us with no clear evidence of a peer effect. Both of these conclusions are in line with panel B of Figure 3. The situation in the END treatment is quite the opposite in that in column 2 of Table 9 we find evidence of a peer effect, with the coefficient for Public Official A's behaviour being significant at the 1 per cent level, while the coefficient for deterrence is not significant.

Table 9: Amount embezzled by Public Official B—by treatment (Tobit)

	Main effects		Interaction only		Full model	
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	ED	END	ED	END	ED	END
A's behaviour (honest=1)	44.371	-446.445***			-104.255	-267.928*
	[101.064]	[88.443]			[202.787]	[143.088]
Detection level	-18.028***	-3.923			-20.312***	-0.535
	[4.713]	[3.881]			[5.442]	[4.429]
A's behaviour=0 # detection level			-19.024***	2.584		
			[4.824]	[4.117]		
A's behaviour=1 # detection level			-14.916**	-25.947***	8.694	-14.718
			[6.201]	[5.892]	[10.289]	[9.325]
Constant	1,915.492***	1,829.088***	1,926.186***	1,733.262***	1,955.257***	1,792.272***
	[149.184]	[149.398]	[145.967]	[148.680]	[156.716]	[151.246]
Sigma_u constant	636.315***	796.107***	636.778***	802.957***	636.972***	797.464***
	[92.486]	[106.146]	[92.542]	[106.895]	[92.572]	[106.260]
Sigma_e constant	1,016.112***	871.252***	1,015.675***	872.039***	1,015.717***	869.630***
	[41.540]	[32.969]	[41.518]	[33.006]	[41.521]	[32.906]
Observations	**640**	**680**	**640**	**680**	**640**	**680**
Subjects	**32**	**34**	**32**	**34**	**32**	**34**

Note: Standard errors in brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: Authors' calculations.

We also see differences between the ED and the END treatments when looking at the effect of detection probability according to Public Official A's type (Equation 2). In the ED treatment, a higher likelihood of detection significantly decreases the amount embezzled by Public Official B on average regardless of whether Public Official A is corrupt or honest (Table 9, column 3). Furthermore, the difference in the coefficients is not significant at conventional levels (p-value = 0.4228, chi2). Table 9, column 4 tells us that in the END treatment the level of detection has no effect on the amount embezzled when that level has been enacted by corrupt Public Official A. However, anti-corruption laws enacted by honest Public Officials A do decrease the amount embezzled by Public Officials B significantly at the 1 per cent level. This difference between the effectiveness of policies promulgated by corrupt and honest Public Official A is significant at the 1 per cent level (p-value = 0.000, chi2). We take this as evidence of a legitimacy effect in the amount embezzled.

The full model (Equation 3) for the ED treatment in Table 9, column 5 shows no significant peer effect but again indicates the presence of a deterrence effect. Table 5, column 5, further confirms that there are no significant differences in the average marginal effects of the type of Public Official A at each detection level. The full model for the END treatment (Table 9, column 6) finds evidence for a peer effect at the 10 per cent level of significance and no evidence of a simple deterrence effect. The interaction term is negative but not statistically significant. Table 5, column 6 shows that changing Public Official A's type from corrupt to honest would lead Public Official B to embezzle a lower amount of funds at all levels of detection.

Externally vs internally imposed detection

We conclude our analysis by again considering if externally imposed detection mechanisms operate differently from their endogenously chosen equivalent. Figure 2.B shows that when the detection probability is 0 per cent, the average amount embezzled by a Public Official B is 1,204ECU in the NoD treatment, 1,786ECU in the ED treatment, and 1,562ECU in the END treatment. The regression results in Table 7, panel B, column 1 indicate that when the detection level is zero, the amount of embezzlement is significantly higher in the ED treatment (at the 1 per cent level) and in the END treatment (at the 5 per cent level), compared to the NoD treatment. There is also evidence of a peer effect when there is no probability of detection as honest behaviour on the part of Public Official A significantly decreases the amount embezzled by Public Official B.

Figure 2.B also shows that when the detection probability is 30 per cent the average amounts embezzled are 1,376ECU, 1,488ECU, and 1,198ECU in the ED, END, and XND treatments respectively. Table 7, panel B, column 3 tells us that when the detection probability is 30 per cent there is no significant difference between the ED and the XND treatments on the one hand, and the END and the XND treatment on the other hand. The peer effect is significant at the 10 per cent level implying that an honest decision made by Public Official A has a socially desirable negative effect on the amount embezzled by Public Official B. We do not see any evidence of an interactive effect of treatment (i.e. institutional framework) and type of Public Official A.

5 Theoretical model

In this section, we consider a variant of the guilt-aversion model proposed by Charness and Dufwenberg (2006) and Battigalli and Dufwenberg (2007) to explain our results. They analyse economic agents who are aversive to acting against the expectations that others have concerning their behaviour. A typical situation where there are strong expectations about behaviour is when there is a social norm that promotes certain behavioural patterns. Thus social norms can be considered a specific application of the guilt-aversion setup. The model of Battigalli and Dufwenberg (2007) belongs to a class of game theory models called psychological games which distinguish themselves from other models in that players' payoffs depend explicitly on beliefs. In what follows, for the sake of simplicity, we will adopt an approach that circumvents this complication but holds true to the main relevant characteristics of the original psychological game model when there is common knowledge of the normative expectations (Miettinen 2013).

In our setting, the normative behaviour requires that the officials not embezzle any funds. If Public Official B deviates from the norm and embezzles $e_B > 0$, then the harm to the recipient organization is $h(e_B) = Ke_B$ (with K equaling 2 in our experimental setup where the transfers are multiplied by two), and the gain to Public Official B from breaching is $g(e_B) = e_B$.

Public Official B is prone to guilt in that if there is a social norm of no embezzlement, he/she feels bad about embezzling only if Public Official A does not embezzle, i.e. $e_A = 0$.[10] If Public Official A embezzles, then Public Official B does not feel bad about embezzling either. Public Official B's intrinsic guilt cost, $\theta_B \gamma(h_B)$, in the case he/she is the only one to embezzle, is continuous, non-negative, and increasing in the inflicted harm with $\gamma'(0) = 0$ and $\gamma' \geq 0$ and $\gamma'' > 0$. For instance, $\gamma(h) = 1/2 \times h^2$ satisfies the assumptions.

Let us denote by U_B Public Official B's expected utility; $e_A \in \{0,1\}$ captures the binary embezzlement choice of Public Official A, with 0 and 1 representing honest and corrupt behaviour, respectively; $e_B \in [0,2280]$ measures the possible range of embezzlement by Public Official B; and p is the detection probability chosen by Public Official A in a preceding stage. By Public Official A's embezzlement behaviour, we can write the difference between Public Official B's utility when embezzling $e_B > 0$ and not embezzling (i.e. Public Official B's incentive to embezzle) as follows:

$$U_B(e_B, e_A; p) - U_B(0, e_A; p) = (1-p)e_B - 1140p - \theta_B (1-e_A)\gamma\,[\mathrm{h}(e_B)]$$

Or

$$U_B(e_B, e_A; p) - U_B(0, e_A; p) = \begin{cases} (1-\mathrm{p})e_B - 1140\mathrm{p} + - \theta_B\, \gamma\big(\mathrm{h}(e_B)\big) & \text{if } e_A = 0 \\ (1-\mathrm{p})e_B - 1140\mathrm{p} & \text{if } e_A = 1 \end{cases} \quad (6)$$

There is an intrinsic cost of violating the social norm of no embezzlement, but only if Public Official A does not embezzle. The parameter θ_B captures Public Official B's susceptibility to guilt. For a given inflicted harm, a Public Official B with higher susceptibility to guilt suffers a higher cost.

The assumptions imply that if Public Official B inflicts strictly positive harm on others, then the guilt cost is strictly positive. Otherwise it is zero. Very small deviations only induce very little guilt while larger deviations matter more; a natural feature not incorporated in constant guilt cost models (Ellingsen and Johannesson 2004; López-pérez 2008, 2012).

Case 1: Let's first consider the case where *Public Official A embezzles some of the funds* and thus transgresses the norm of no embezzlement, i.e. the latter case in Equation (6). In this case, Public Official B does not feel bad about breaking it too and thus Public Official B engages in embezzlement if:

$$U_B(e_B, 1; p) - U_B(0, 1; p) = (1-p)e_B - 1140p > 0$$

Thus when Public Official A embezzles, Public Official B's incentive to embezzle is decreasing in p.[11] Notice, however, that the incentive to embezzle is also increasing in e_B so Public Official B prefers embezzling all the funds (i.e. $e_B^* = 2280$) if she prefers embezzling at all. Inserting $e_B^* = 2280$ and simplifying yields a condition for optimality of full embezzlement: $2 > p\,/\,(1-p)$, which holds true when p ≤ 1/3. Thus in the context of our experiment, *the theory*

---

[10] This is in line with other models (Bicchieri 2005; Lopez-Perez 2008), casual observations (see Bicchieri 2005) and experimental evidence (Miettinen and Suetens 2008).

[11] In a logit-quantal response equilibrium (QRE) (Goeree et al. 2010), the embezzlement would have full support (i.e. all actions from 0 to 2280 would be chosen with a positive probability). The fact that the incentive to embezzle is decreasing in p would then imply that the probability of embezzlement is decreasing in p in a QRE framework.

predicts that if Public Official A embezzles, Public Official B always prefers embezzling, but that the incentive to embezzle is decreasing in p.

Case 2: Suppose now that *Public Official A does not embezzle any funds,* i.e. the first row of (6). In this case, the optimality condition for an Official B is given by:

$$U_B(e_B, 0; p) - U_2(0, 0; p) = (1-p)\, e_B - 1140p - \theta_B\, \gamma(2 \times e_B) > 0,$$

which gives the necessary and sufficient condition for embezzlement by Public Official B.[12] The incentive to embezzle is again decreasing in p. By embezzling, Public Official B gains e_B but runs the risk of losing both e_B and the flat salary with probability p. The term $-\theta_B\, \gamma(K \times e_B)$ measures guilt about transgressing the social norm where γ is a convex function and $K \times e_B$ is the extent of harm inflicted on the recipient organization. The parameter θ_B captures Public Official B's susceptibility to guilt and $\theta_B = 0$ correponds to a standard self-regarding Public Official B. Guilt is an increasing and convex function of the inflicted harm. The optimal level of embezzlement by Public Official B is given implicitly by $(1-p) = \theta\, \gamma'(K \times e_B^{*,\theta})$ where $e_B^{*,\theta}$ is decreasing in p is provided that $\theta_B > 0$. In the special case where $\gamma(h_B) = 1/2 \times h_B^2$, we have $e_B^* = (1-p)/K^2\, \theta_B$, so that as θ_B tends towards 0 (the standard type), Public Official B embezzles all funds. When θ_B tends to infinity, Public Official B embezzles nothing.

Let us summarize our findings in the following proposition.

Proposition. *Suppose Public Official B is guilt-averse:*

- *If Public Official A embezzles funds, then a guilt-averse Public Official B always embezzles funds and chooses* $e_B^* = 2280$. *The embezzled amount is independent of p.*
- *If Public Official A does not embezzle any funds, then a guilt-averse Public Official B prefers to embezzle the amount implicitly given by* $(1-p) = \theta\, \gamma'(2 \times e_B^{*,\theta})$. *The amount* $e_B^{*,\theta}$ *is decreasing in p and in θ.*
- *If Public Official A embezzles, then the probability of detection is ineffective in curbing the embezzlement of a guilt-averse Public Official B. Yet, if Public Official A does not embezzle any funds, then the p is effective in curbing the embezzlement of a guilt-averse Public Official B. Thus, there is a legitimacy effect on Public Official B's behaviour.*

Thus our model predicts that for a decision maker who is guilt-averse, there is both a peer effect (Public Official A's embezzlement triggers more embezzlement by Public Official B) and a legitimacy effect (Public Official A's embezzlement renders deterrence less effective). More generally, let us denote the optimal embezzlement level of Public Official B by $e_B^*(e_A, p)$. The deterrence effect conditional on (e_A, p) is formally defined by $\frac{\partial e_B^*(e_A,p)}{\partial p}$ and the unconditional deterrence effect is $\int \frac{\partial e_B^*(v,p)}{\partial p} dv$. The peer effect at (e_A, p) is defined by $\frac{\partial e_B^*(e_A,p)}{\partial e_A}$ or in our discrete Public Official A-behaviour case by $e_B^*(0,p) - e_B^*(1,p)$ and the unconditional peer effect is equal to $\int e_B^*(0,s) - e_B^*(1,s) ds$. The legitimacy effect is then given by

[12] We assume that an indifferent Public Official B does not embezzle funds thus breaking indifference in favour of norm abidance. See for instance Demichelis and Weibull (2008).

$\frac{\partial}{\partial p}(e_B^*(0,p) - e_B^*(1,p))$, or alternatively, $\frac{\partial e_B^*(0,p)}{\partial p} - \frac{\partial e_B^*(1,p)}{\partial p}$, so that the effect of p on e_B^* depends on whether Public Official A is honest or corrupt.

Note also that the incentive to embezzle may be decreasing in p, for instance, although the optimal action is unaffected by p. This is the case for example if $U_B(e_B\,, e_A; p) - U_B(0\,, e_A; p)$ is positive for all p and for all e_i but the difference is still decreasing in p for all e_i.

The theoretical predictions are in line with the observed patterns in the END treatment where the deterrence by a corrupt Public Official A is ineffective but an honest Public Official A can effectively reduce embezzlement by choosing a higher level of detection. Relative to the ED treatment, one might conjecture that institutions which treat both parties unequally and unsymmetrically are less susceptible to trigger a shared perception of normative behaviour where violations of the norm by Public Official A crowd out the intrinsic motivation by Public Official B. The absence of a norm which Public Official B can feel bad about breaking can be captured in Case 1. Mainly Public Official B's incentive to embezzle will be decreasing in p but unaffected by e_A. As a result, in the treatment ED, only the pecuniary deterrence effect is observed but the norm-motivated effects of peer influence and legitimacy are absent.

6 Concluding remarks

This paper draws on data obtained from a framed laboratory experiment carried out in Kenya to examine the roles of peer effects, deterrence effects, and legitimacy effects in the fight against corruption. In our regression analysis, we labelled the main effect of Public Official A's behaviour as a peer effect, and that of the detection level as a deterrence effect. The legitimacy effect refers to the interaction between A's behaviour and detection levels. This captured the idea that deterrence can be less effective when chosen by a corrupt A.

Crucially, we found that the importance of these effects depended on the institutional framework in which our 'public officials' found themselves operating. When policy makers are exempt from their own laws we find that a strong deterrence effect, a greater chance of being detected and punished, reduces the likelihood and the extent of corruption. This effect does not depend on the behaviour of the policy maker. In settings in which equality before the law is observed and policy makers are liable to be caught in their own net, we find that detection policies are only an effective deterrent when promulgated by honest policy makers. This legitimacy effect is evident alongside a simple peer effect. We also found that externally imposed rules may be superior to equally stringent rules originating from a corrupt internal policy maker. Once again, this existence of this effect was dependent on the internal policy maker being subject to the provisions of the policy.

Our findings offer several important implications in fighting corruption for policy makers and other interested parties—subject to the usual external validity caveats of experimental economics which we will address briefly below. Firstly, our results add further evidence as to the potential for detection and punishment mechanisms to play a role in curbing corruption. Our findings of a peer effect suggest that creating a culture of honesty among the top-rank officials in systems such as the one in our experiment can have knock-on, or perhaps trickle-down, effects on others within the organization or society (Moxnes and Van det Heijden 2003; Güth et al. 2007; Levati et al. 2007; Cappelen et al. 2015). Our finding of a strong legitimacy effect adds more weight to this argument in that fostering such an honest ethic may result in the same policy being more effective. Moreover, internally generated anti-corruption detection mechanisms will only be effective in institutional settings with equality before the law when the policy maker is honest. If this condition is not met, exogenously imposed rules are preferable.

The results on the effects of institutional settings on legitimacy can be related to the perceived procedural fairness of the system. Indeed, the procedural fairness literature suggests that legitimacy springs from a shared perception between all relevant parties and outsiders about the fairness of the procedures applied—though outcomes may be unequal, at least everyone acts under a common set of rules that equally apply to all (Lind 2001; Tyler 2004). Perceived procedural fairness promotes compliance with the verdicts of the authority. Since the seminal work of Thibaut and Walker (1975), various studies have come to establish and support these views (see e.g. Lind 2001; Falk et al. 2003; Tyler 2004; Bolton et al. 2005). Therefore, it is somewhat surprising that in our setting, the asymmetric rules of the game promote compliance independent of the behaviour of the authority that decides upon the anti-corruptive measures. Another potential explanation may relate to risk-taking behaviour in-group vs individually. There are a number of experiments that suggest higher risk-taking behaviour in groups compared to isolated individuals (for example, Yechiam et al. 2008; Cooper and Rege 2011; Bougheas et al. 2013; Lahno and Serra-Garcia 2015). As detection applies to both officials in the END treatment, this may create a group effect when Public Official A decides to embezzle despite setting a positive detection level and despite the fact that in our detection mechanism, independent draws are carried out for Officials A and B. In the ED treatment, no such effect exists as deterrence applies only to Public Official B. Finally, perhaps the explanation for our results lies in the fact that the ultimate authority that chooses the symmetric or asymmetric rules is not the first-moving official but the experimenter. Given the importance of the institutional framework in our results, future research could usefully endogonize the choice of whether the fight against corruption applies to all levels of the society or organization or whether immunity is granted to the top level which has an influence on the anti-corruption institutions.

As to how much faith one can take in the applicability of lab results to the 'real world', a few arguments can be noted. Firstly, while the magnitude of any given effect may not carry over from the lab to the field, for an experiment to be useful we need only qualitative external validity (Camerer 2014; Kessler and Vesterlund 2014). Secondly, external validity problems are not unique to experimental results (Falk and Heckman 2009; Kessler and Vesterlund 2014). Finally, the very nature of corruption makes data difficult to collect and means that there are practical and ethical challenges to evaluating anti-corruption policies and institutions in a field setting (Plott 1999; Klemperer 2004; Milgrom 2004; Armantier and Boly 2012). Laboratory studies allow us to get a handle on what might work in a low-cost and ethically feasible setting (Dusek et al. 2005; Abbink 2006).

References

Abbink, K. (2006). 'Laboratory Experiments on Corruption'. In S. Rose-Ackerman (ed.), *The Handbook of Corruption*. Cheltenham, UK, and Northampton, US: Edward Elgar Publishers.

Abbink, K., B. Irlenbusch, and E. Renner (2002). 'An Experimental Bribery Game'. *Journal of Law, Economics, and Organization*, 18(2): 428–54.

Abbink, K., and D. Serra (2012). 'Anticorruption Policies: Lessons from the Lab'. *Research in Experimental Economics*, 15: 77–115.

Akerlof, R. (2016). 'The Importance of Legitimacy'. *The World Bank Economic Review*. doi: 10.1093/wber/lhw009.

Armantier, O., and A. Boly (2012). 'On the External Validity of Laboratory Experiments on Corruption'. In D. Serra and L. Wantchekon (eds), *New Advances in Experimental Research on Corruption (Research in Experimental Economics)*, 15: 117–44. Bingley: Emerald Group Publishing Limited.

Azfar, O., and W.R. Nelson Jr (2007). 'Transparency, Wages, and the Separation of Powers: An Experimental Analysis of Corruption'. *Public Choice*, 130(3–4): 471–93.

Battigalli, P., and M. Dufwenberg (2007). 'Guilt in Games'. *The American Economic Review,* 97(2): 170–76.

Bandiera, O., I. Barankay, and I. Rasul (2010). 'Social Incentives in the Workplace'. *Review of Economic Studies*, 77(2): 417–58.

Bardsley, N., and R. Sausgruber (2005). 'Conformity and Reciprocity in Public Good Provision'. *Journal of Economic Psychology*, 26(5): 664–81.

Barr, A., M. Lindelow, and P. Serneels (2009). 'Corruption in Public Service Delivery: An Experimental Analysis'. *Journal of Economic Behavior & Organization*, 72(1): 225–39.

Basu, K. (2015). 'The Republic of Beliefs: A New Approach to Law and Economics'. World Bank Policy Research Working Paper 7259. Washington, DC: World Bank.

Beams, J.D., R.M. Brown, and L.N. Killough (2003). 'An Experiment Testing the Determinants of Non-compliance with Insider Trading Laws'. *Journal of Business Ethics*, 45(4): 309–23.

Becker, G.S. (1968). 'Crime and Punishment: An Economic Approach'. *Journal of Political Economy*, 76(2): 169–217.

Bicchieri, C. (2005). *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge: Cambridge University Press.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1998). 'Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades'. *The Journal of Economic Perspectives*, 12(3): 151–70.

Bolton, G.E., J. Brandts, and A. Ockenfels (2005). 'Fair Procedures: Evidence from Games Involving Lotteries'. *Economic Journal*, 115(506): 1054–76.

Boly, A., and R. Gillanders (2016). 'Anti-corruption Policy-making, Discretionary Power, and Institutional Quality'. WIDER Working Paper 2016/88. Helsinki: UNU-WIDER.

Bougheas, S., J. Nieboer, and M. Sefton (2013). 'Risk-taking in Social Settings: Group and Peer Effects'. *Journal of Economic Behavior & Organization*, 92: 273–83.

Brandts, J., and A. Schram (2001). 'Cooperation and Noise in Public Goods Experiments: Applying the Contribution Function Approach'. *Journal of Public Economics*, 79(2): 399–427.

Camerer, C.F. (2014). 'The Promise and Success of Lab-Field Generalizability in Experimental Economics: A Reply to Levitt and List'. In G. Fréchette and A. Schotter (eds), *The Methods of Modern Experimental Economics*. Oxford: Oxford University Press.

Cappelen, A.W., B.A. Reme, E.Ø. Sørensen, and B. Tungodden (2015). 'Leadership and Incentives'. *Management Science*, Forthcoming.

Charness, G., and M. Dufwenberg (2006). 'Promises and Partnership'. *Econometrica*, 74(6): 1579–601.

Chen, D. (2013). 'The Deterrent Effect of the Death Penalty? Evidence from British Commutations during World War I'. NBER Working Paper. Cambridge, MA: National Bureau of Economic Research.

Cooper, D.J., and M. Rege (2011). 'Misery Loves Company: Social Regret and Social Interaction Effects in Choices under Risk and Uncertainty'. *Games and Economic Behavior*, 73(1): 91–110.

d'Adda, G., D. Darai, and R.A. Weber (2014). 'Do Leaders Affect Ethical Conduct?'. University of Zurich Working Paper 167. Zurich: University of Zurich.

Demichelis, S., and J.W. Weibull (2008). 'Language, Meaning, and Games: A Model of Communication, Coordination, and Evolution'. *The American Economic Review*, 98(4): 1292–311.

Dineen, B.R., R.J. Lewicki, and E.C. Tomlinson (2006). 'Supervisory Guidance and Behavioral Integrity: Relationships with Employee Citizenship and Deviant Behavior'. *Journal of Applied Psychology*, 91(3): 622–35.

Dusek, L., A. Ortmann, and L. Lizal (2005). 'Understanding Corruption and Corruptibility through Experiments: A Primer'. *Prague Economic Papers*, 14(2): 147–62.

Dong, B., U. Dulleck, and B. Torgler (2012). 'Conditional Corruption'. *Journal of Economic Psychology*, 33(3): 609–27.

Eide, E., P.H. Rubin, and J.M. Shepherd (2006). *Economics of Crime*. Boston, MA: Now Publishers Inc.

Ellingsen, T., and M. Johannesson (2004). 'Promises, Threats and Fairness'. *The Economic Journal*, 114(495): 397–420.

Falk, A., E. Fehr, and U. Fischbacher (2003). 'On the Nature of Fair Behavior'. *Economic Inquiry*, 41(1): 20–26.

Falk, A., and J.J. Heckman, (2009). 'Lab Experiments Are a Major Source of Knowledge in the Social Sciences'. *Science*, 326: 535–538.

Falk, A., and A. Ichino (2006). 'Clean Evidence on Peer Effects'. *Journal of Labor Economics*, 24(1): 39–57.

Falk, A., and M. Kosfeld (2006). 'The Hidden Costs of Control'. *American Economic Review*, 96(5): 1611–30.

Fehr, E., and B. Rockenbach (2003). 'Detrimental Effects of Sanctions on Human Altruism'. *Nature*, 422(6928): 137–140.

Fisman, R., and J. Svensson (2007). 'Are Corruption and Taxation Really Harmful to Growth? Firm Level Evidence'. *Journal of Development Economics*, 83(1): 63–75.

Fischbacher, U., and S. Gächter (2010). 'Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments'. *American Economic Review*, 100(1): 541–56.

Fortin, B., G. Lacroix, and M.C. Villeval (2007). 'Tax Evasion and Social Interactions *Journal of Public Economics*, 91(11): 2089–112.

Francois, P. (2000). '"Public Service Motivation" as an Argument for Government Provision'. *Journal of Public Economics*, 78(3): 275–99.

Francois, P. (2003). 'Not-For-Profit Provision of Public Services'. *Economic Journal*, 113(486): C53-C61.

Gatti, R., S. Paternostro, and J. Rigolini (2003). 'Individual Attitudes Toward Corruption: Do Social Effects Matter?'. *World Bank Policy Research Working Paper*, (3122). Washington, DC: World Bank.

Goeree, J.K., C.A. Holt, and T.R. Palfrey (2010). 'Quantal Response Equilibria'. In S. Durlauf and L. Blume (eds), *Behavioural and Experimental Economics*. London: Palgrave Macmillan.

Gupta, S., H. Davoodi, and R. Alonso-Terme (2002). 'Does Corruption Affect Income Inequality and Poverty?'. *Economics of Governance*, 3(1): 23–45.

Güth, W., M.V. Levati, M. Sutter, and E. Van der Heijden (2007). 'Leading by Example with and without Exclusion Power in Voluntary Contribution Experiments'. *Journal of Public Economics*, 91(5): 1023–42.

Hanna, R., S. Bishop, S. Nadel, G. Scheffler, and K. Durlacher (2011). 'The Effectiveness of Anti-corruption Policy'. EPPI Centre Report (No. 1909). London: EPPI-Centre.

Jones, G.E., and M.J. Kavanagh (1996). 'An Experimental Examination of the Effects of Individual and Situational Factors on Unethical Behavioral Intentions in the Workplace'. *Journal of Business Ethics*, 15(5): 511–23.

Kagan, R.A., and J.T. Scholz (1984). 'The Criminology of the Corporation and Regulatory Enforcement Strategies'. In K. Hawkins and J.M. Thomas (eds), *Enforcing Regulation*. Boston: Kluwer-Nijhoff.

Kessler, J.B., and L. Vesterlund (2014). 'The External Validity of Laboratory Experiments: Qualitative rather than Quantitative Effects'. In G. Fréchette and A. Schotter (eds), *The Methods of Modern Experimental Economics*: Oxford: Oxford University Press.

Keizer, K., S. Lindenberg, and L. Steg (2008). 'The Spreading of Disorder'. *Science*, 322(5908): 1681–85.

Klemperer, P. (2004). *Auctions: Theory and Practice*: Princeton: Princeton University Press.

Kornhauser, L.A. (1984). 'The Great Image of Authority'. *Stanford Law Review*, 349–89.

Lambsdorff, J.G. (2015). 'Preventing Corruption by Promoting Trust: Insights from Behavioral Science'. 69(15) Passauer Diskussionspapiere: Volkswirtschaftliche Reihe. Passau: Universität Passau.

Lahno, A.M., and M. Serra-Garcia (2015). 'Peer Effects in Risk Taking: Envy or Conformity?'. *Journal of Risk and Uncertainty*, 50(1): 73–95.

Lefebvre, M., P. Pestieau, A. Riedl, and M.C. Villeval (2015). 'Tax Evasion and Social Information: An Experiment in Belgium, France, and the Netherlands'. *International Tax and Public Finance*, 22(3): 401–25.

Levati, M.V., M. Sutter, and E. Van der Heijden (2007). 'Leading by Example in a Public Goods Experiment with Heterogeneity and Incomplete Information'. *Journal of Conflict Resolution*, 51(5): 793–818.

Levi, M., and A. Sacks (2009). 'Legitimating Beliefs: Sources and Indicators'. *Regulation & Governance*, 3(4): 311–33.

Lind, E.A. (2001). 'Fairness Heuristic Theory: Justice Judgments as Pivotal Cognitions in Organizational Relations'. *Advances in Organizational Justice*, 56: 88.

López-Pérez, R. (2008). 'Aversion to Norm-breaking: A Model'. *Games and Economic Behavior*, 64(1): 237–67.

López-Pérez, R. (2012). 'The Power of Words: a Model of Honesty and Fairness'. *Journal of Economic Psychology*, 33(3): 642–58.

Mauro, P. (1995). 'Corruption and Growth'. *Quarterly Journal of Economics*, 681–712.

Miettinen, T. (2013). 'Promises and Conventions—an Approach to Pre-play Agreements'. *Games and Economic Behavior*, 80: 68–84.

Miettinen, T., and S. Suetens (2008). 'Communication and Guilt in a Prisoner's Dilemma'. *Journal of Conflict Resolution*, 52(6): 945–60.

Milgrom, P. (2004). *Putting Auction Theory to Work*. Cambridge: Cambridge University Press.

Moxnes, E., and E. Van der Heijden (2003). 'The Effect of Leadership in a Public Bad Experiment'. *Journal of Conflict Resolution*, 47(6): 773–95.

Olken, B.A. (2007). 'Monitoring Corruption: Evidence from a Field Experiment in Indonesia'. *Journal of Political Economy*, 115(2).

Papachristos, A.V., T.L. Meares, and J. Fagan (2012). 'Why Do Criminals Obey the Law? The Influence of Legitimacy and Social Networks on Active Gun Offenders'. *The Journal of Criminal Law and Criminology*, 397–440.

Paternoster, R. (1987). 'The Deterrent Effect of the Perceived Certainty and Severity of Punishment: A Review of the Evidence and Issues'. *Justice Quarterly*, 4(2): 173–217.

Pierce, L., and J. Snyder (2008). 'Ethical Spillovers in Firms: Evidence from Vehicle Emissions Testing'. *Management Science*, 54(11): 1891–903.

Plott, C.R. (1999). 'Policy and the Use of Laboratory Experimental Methodology in Economics'. In L. Luini (ed.), *Uncertain Decisions Bridging Theory and Experiments*. Boston: Kluwer Academic Publishers.

Sacerdote, B. (2011). 'Peer Effects in Education: How Might they Work, How Big Are they and How Much Do we Know Thus Far?'. *Handbook of the Economics of Education*, 3: 249–77.

Schnellenbach, J. (2007). 'Tax Morale and the *Legitimacy* of Economic Policy'. *Homo Oeconomicus*, 24(1): 21–45.

Schulze, G.G., and B. Frank (2003). 'Deterrence versus Intrinsic Motivation: Experimental Evidence on the Determinants of Corruptibility'. *Economics of Governance*, 4(2): 143–60.

Serra, D. (2012). 'Combining Top-down and Bottom-up Accountability: Evidence from a Bribery Experiment'. *Journal of Law, Economics, and Organization*, 28(3): 569–87.

Shang, J., and R. Croson (2009). 'A Field Experiment in Charitable Contribution: The Impact of Social Information on the Voluntary Provision of Public Goods'. *Economic Journal*, 119(540): 1422–39.

Smith, S., F. Windmeijer, and E. Wright (2015). 'Peer Effects in Charitable Giving: Evidence from the (Running) Field'. *Economic Journal*, 125(585): 1053–71.

Thibault, J., and L. Walker (1975). *Procedural Justice: A Psychological Analysis*. New York: L. Erlbaum Associates.

Tyler, T.R. (1990). *Why People Obey the Law*. Princeton and Oxford: Princeton University Press.

Tyler, T.R. (2004). 'Enhancing Police Legitimacy'. *The Annals of the American Academy of Political and Social Science*, 593(1): 84–99.

Tyler, T.R. (2006). 'Psychological Perspectives on Legitimacy and Legitimation'. *Annual Review of Psychology*, 57: 375–400.

Weber, M. (1964). *The Theory of Social and Economic Organization*. Talcott Parsons (ed.). New York: Free Press.

Yechiam, E., M. Druyan, and E. Ert (2008). 'Observing Others' Behavior and Risk Taking in Decisions from Experience'. *Judgment and Decision Making*, 3(7): 493–500.