Organizational cultures of corruption*

Patrick Schneider and Gautam Bose †

September 7, 2015

Abstract

Systematic differences in the incidence of corruption between countries can be explained by models of coordination failure that suggest that corruption can only be reduced by a "big push" across an entire economy. However, there is significant evidence that corruption is often sustained as an organizational culture, and can be combated with targeted effort in individual organizations one at a time. In this paper we propose a model that reconciles these two theories of corruption.

We explore a model of corruption with two principal elements. First, agents suffer a moral cost if their corruption behavior diverges from the level they perceive to be the social norm; second, the perception of the norm is imperfect; it gives more weight to the behavior of colleagues with whom the agent interacts regularly. This leads to the possibility that different organizations within the same country may stabilize at widely different levels of corruption. Further, the level of corruption in an organization is persistent, implying that some organizations may have established internal 'cultures' of corruption. The organizational foci are determined primarily by the opportunities and (moral) costs of corruption. Depending on the values of these parameters the degree of corruption across departments may be relatively uniform or widely dispersed.

These results also explain another surprising empirical observation: that in different countries similar government departments such as tax and education rank very differently relative to each other in the extent to which they are corrupt. This is difficult to explain in incentive-based models if similar departments face similar incentives in different countries.

Keywords: corruption, cognitive dissonance, perception, social norms.

*This paper is partly based on Schneider's undergraduate honors thesis submitted to UNSW. We are indebted to Zhanar Akhmetova, Ariel Ben Yishay, Richard Holden, Hodaka Morita and Alberto Motta for comments at various stages of this project. Correspondence to g.bose@unsw.edu.au.

†University of New South Wales.

1. Introduction

Levels of corruption differ significantly between countries.[1] Coordination models provide a standard way to explain such differences and suggest that anti-corruption efforts should involve 'big-push' policies that ensure all agents shift expectations and behavior simultaneously to the desirable equilibrium (see, e.g., Andvig and Moene (1990) and the references therein, Murphy et al. (1993); Nabin and Bose (2008)). The persistence and pervasiveness of corruption in some countries coupled with its systematic absence in others strongly support the insights of coordination models.

Corrupt practices also vary substantially across different regions or organizations within individual countries. This within-country variation can be explained at the level of organizations, with differences in corruption caused by differing incentives or management regimes and exacerbated by selection.[2] These lower level models suggest corruption can be cleaned up within organizations by altering the incentive structures that agents face. There are examples in the literature of successful anti-corruption efforts targeting individual organizations, for example in Klitgaard (1988).

There is a disjuncture in the insights that these approaches provide. The latter suggests that it is feasible to clean up organizations with targeted policies. The former approach, on the other hand, suggests that this should not be possible—only an economy-wide effort will work. However, while they produce conflicting policy recommendations, both approaches have empirical support.

In this paper we propose a simple model of coordination on social norms (Rabin, 1994; Fischer and Huddart, 2008) to reconcile the two approaches. An agent in our model chooses how much to exploit opportunities for corruption but incurs a cost in doing so. The magnitude of this cost depends on his individual perception of the social norm regarding 'acceptable' corruption, which is an aggregation of the positions held by his peers within the same organization as well as agents across the society. It is crucial to our model that the agent's perception is more heavily influenced by the stances of his immediate peers than by those outside his organization. In equilibrium, each organization coordinates on a local norm which we term its 'organizational focus'. Different organizations may coordinate on widely different foci even when agents are homogeneous, self-selection is absent, and there are no differences in incentives. This result is consistent with the empirical observation that corruption levels between organizations vary more than differences in incentives can explain (Schneider, 2011).

We define the 'social norm' as the employment-weighted average behavior across all organizations. The individual agent's *perception* of the social norm, however, is imperfect; it

[1] Evident in casual observation as well as increasing numbers of corruption measures such as Transparency International's indices.

[2] An important institutionalized selection mechanism that has been long known in the Indian context is the sale of "lucrative" administrative positions based on the bribe-income that an incumbent is able to earn during the tenure of his posting. See Wade (1982, especially pp. 302-307) for an early and detailed study, and also Wade (1985).

accords positive weight to the behavior of his organizational colleagues (the local norm) as well as the true social norm. When his perception is limited only to the local norm, the culture in each organization is independent of the rest of the economy. Similarly, when agents accurately perceive the social norm, all organizations in the economy share a homogeneous culture. However, when agents are influenced by both their local and the global norm, differences between organizations can be sustained in equilibrium. We identify conditions under which the equilibrium distribution of organizational foci in an economy is non-degenerate.

An equilibrium in our model is a state in which each agent in each organization optimally responds to his perceived social norm, and these responses give rise to the corresponding organizational foci. Note that it is not necessary that all organizations have the same focus in equilibrium. There may be multiple sustainable foci within an equilibrium. Further note that multiple equilibria can obtain, each with its own distribution of organizational foci, depending on the extant global norm. For example, for a given global norm, there may be an equilibrium in which some organizations focus on high and some on low levels of corruption, and another equilibrium in which all organizations focus on the same level of corruption. The requisite global norm is maintained in equilibrium if appropriate proportions of organizations occupy the admissible organizational foci.

Even with homogeneous agents and exogenously determined incentives, our model leads to a richer set of policy possibilities than pure coordination models. Since organizational foci are determined in part by endogenously generated norms, the extent of corruption within an organization can be successfully reduced by a targeted initiative even without changes in the rest of the economy. However, the extent of corruption within the organization also does contribute to the social average. As a result, changes in one organization will flow through to others, at least in small measure. We know that such targeted efforts do yield results, as documented, for example, by Klitgaard (1988) in the case of the Hong Kong police force and of the Philippines income tax department. An effort to rid a corrupt economy of the malady could thus target organizations sequentially, or it could target them simultaneously as in a "big push". If a sequential strategy is followed, then the most efficient sequencing is to first target the organizations that will have the largest impact on the social norm per unit of cost.

The mechanism we employ utilizes a social dynamic that is well-established in social psychology and other related literatures. People experience costs in terms of social opprobrium or negative feelings when they do something perceived to be 'bad' (Coleman, 1990; Basu, 2000; Akerlof, 1976; Elster, 1989). But these costs depend on the attitudes of those around them because their peers help them get a view of what is normal (Bardhan, 1997; Sardan, 1999; Arendt, 1977), help them develop rationalizations to avoid classifying the activity as bad at all, thus avoiding the cost (Cressey, 1971; Laub, 2006), or involve them in a web of complicity so as to nullify objection (Ashforth and Anand, 2003).[3]

[3]Rasmusen (1996) shows that when the crime rate is high workers with prior criminal convictions face less social stigma (in terms of lower wages). However, the result is driven by adverse selection rather than peer effects: criminals are less productive than non-criminals and in high-crime environments there is a higher proportion of unconvicted criminals, bringing the expected productivity of this population closer to that of

As with the rose, our mechanism by any other name would work as well. Our analytical results derive from two critical assumptions: first, that the individual's perception of the social norm is a convex combination of the true social norm and the local norm, and secondly, that there is a cost when the individual's behavior deviates from this perceived norm. We motivate our model with reference to culture and norm-conforming preferences because we believe this reflects actual behavior. However, the model is capable of other interpretations; for example, the probability of detection, censure and punishment for corrupt acts (as well as praise or reward for eschewing them) may depend on the extent to which an individual's practice deviates from those of his organizational colleagues.

We interpret culture as the modes of behavior and language employed by groups of closely interacting people. There are differences in corruption and attitudes toward it that cannot reasonably be explained with reference to normal incentives. For example, four of the last eight governors of Illinois[4] have criminal convictions, whereas none in Indiana do. Certainly there are differences in the neighboring states but the contrast is stark, and there is no reason to think the incentives are *that* different. Similarly, in their study on shirking in branches of the same corporations located in different regions of Italy, Ichino and Maggi (2000) find systematic differences between northern and southern branches. Further, they find that employees that relocate from one region to the other tend to change their behavior to conform to the local norm.

In the context of comparative corruption between countries, Schneider (2011, Chapter 7), finds unexplained variation in corruption levels between government departments within countries. Not surprisingly, corruption is higher across the board in some countries than in others, and some departments or authorities are on average more corrupt than other departments or authorities once country effects are accounted for. However, the latter effect is significantly less systematic than one would expect on the basis of a simple, incentive-driven model. There is a sufficiently large number of instances in which department A is significantly more corrupt than department B in some countries, while the opposite is true elsewhere. An explanation for this variation is that individual departments idiosyncratically develop distinct cultures of corruption, which is the way we prefer to interpret our results.

Past papers have employed aspects of these dynamics to model corrupt or otherwise 'bad' behavior. The cost of immorality, for example, is introduced exogenously by Blackburn et al. (2006) and Bose and Gangopadhyay (2009) by incorporating corruptible and incorruptible agents, and by Klitgaard (1988) who introduces agents who suffer a fixed moral cost for their actions. Alternately, cognitive dissonance and rationalization has been invoked by Akerlof and Dickens (1982) and Rabin (1994) to show how the interaction between one's behavior, one's own moral line and the behavior of peers can interact to support damaging behavior.

In Glaeser et al. (1996), agents choose whether or not to commit crimes. Some agents are influenced by their neighbors' choices while other agents are not. The latter type induce boundaries that contain the diffusion of influence, forming distinct neighborhoods across

convicted criminals.

[4] Rod Blagojevich (2003-09), George H. Ryan (1999-2003), Dan Walker (1973-77) and Otto Kerner, Jr (1962-68).

which crime rates can vary significantly. Our model is somewhat similar, but instead of different types of agents we rely on organizational boundaries and perception to generate local differences. Fischer and Huddart (2008) develop a model that brings these issues together in a fashion similar to ours—agents choose the level at which they undertake an action and suffer a cost that depends on the (endogenous) social norms among their peers. They find that, within this structure, norms influence the effectiveness of other incentives, determine the ideal way to group agents and also guide self-selection. These two papers come the closest to fully endogenizing social norms, though they rely on heterogeneous agents to produce the key results.

Our model builds on this literature. We introduce locally biased perception to endogenize social norms. This modification allows us to construct a particularly simple model that generates novel insights. The model focuses solely on a 'bad' action and an associated internalized cost that depends on the choices of others. Structured thus, it could apply equally to different settings (e.g. government or private) and to applications other than corruption. The coordination that the model produces follows naturally from its structure. Its utility is in its ability to produce, in an extremely simple framework, a variety of observed behaviors by homogeneous agents that coexist in equilibrium.

We do not invoke heterogeneity between agents or between the incentives for corruption in different organizations. The heuristic discussion of the general model in Section 2 indicates that a large number of foci are possible within a single equilibrium depending on how perceptions are formed. In the closed-form model of Sections 3 and 4 we use a strictly convex moral cost function and a (weakly) concave utility function, and hence obtain equilibria with at most two distinct organizational foci. It will be readily obvious that if agents were heterogeneous, or they were to self-select into organizations based on their own predispositions towards being corrupt, this would strengthen our results and produce a richer spread of organizational foci without contradicting our policy conclusions.

Our emphasis is on the role of social norms in determining corruption, so for clarity we abstract from the possibility that corrupt acts may attract a probabilistic penalty. Alternatively, the cost in the model may be interpreted as already incorporating the expected exogenous penalty.

Section 2 describes the model. Section 3 discusses how each organization's focus is determined. Section 4 establishes the equilibria for the whole economy. Section 5 discusses the policy implications and concludes the paper.

2. The model

2.1. Organizations, agents and perceptions

The economy consists of J departments, indexed by j. Department j employs a number N_j of officials indexed by ij. All officials earn a common wage which we normalize to zero. Each official also faces opportunities for corruption from which he can earn a maximum additional income b_{ij}. The extent to which he actually indulges in these opportunities

$\psi_{ij} \in [0, 1]$ is his choice. His income is then $\psi_{ij} b_{ij}$, from which he derives a benefit

$$U(\psi_{ij} b_{ij}) \quad \text{where} \quad U' > 0 \quad \text{and} \quad U'' \leq 0.$$

While the official can choose his behavior, it is costly to act in a way that deviates from what he perceives to be the social norm ϕ (to be explained below). If the official wishes to maintain behavior $\psi \neq \phi$, he incurs a (psychological or moral) cost:

$$C(\psi, \phi), \quad \text{where} \quad C_\psi > 0,\ C_\phi < 0,\ C_{\psi\psi} > 0 \quad \text{and} \quad C_{\psi\phi} < 0.$$

The cost is increasing and convex in ψ, and decreasing in the perceived level of corruption in the society. Further, the marginal cost of a small increase in personal corruption is smaller, *ceteris paribus*, when the perceived social corruption is larger.

The official does not perfectly observe the norm of his broader society, $\theta \in [0, 1]$. Instead, he forms a perception based on the average of the behavior of those in his department,

$$\psi_j = \frac{\sum_i \psi_{ij}}{N_j} \tag{1}$$

and the society-wide norm θ.We will discuss the genesis of θ later; for the moment we assume it to be exogenous. So the officials in department j perceive the social norm to be[5]

$$\phi_j = \Phi(\psi_j, \theta), \quad \min\{\psi_j, \theta\} \leq \Phi(\psi_j, \theta) \leq \max\{\psi_j, \theta\}, \quad 0 < \Phi_\psi, \Phi_\theta < 1. \tag{2}$$

Each official's payoff is the sum of the utility from his corruption income and his moral cost. The official chooses ψ_{ij} to maximize this objective function, with his perception of the social norm ϕ_j taken as given. Thus his optimization problem is:

$$\max_{\psi_{ij}} \quad V_{ij}(\psi_{ij}, \phi_j) = U(\psi_{ij} b_{ij}) - C(\psi_{ij}, \phi_j) \tag{3}$$

It follows that ψ_{ij} is a function of b_{ij} and ϕ_j. The organization's average behavior is then $\Psi_j(\phi_j, \theta)$ as in (1), where the individual behavior choices ψ_{ij} solve (3).

It is intuitively obvious that, for given θ, an average behavior ψ_j can persist in department j only if it gives rise to a perception ϕ_j which in turn leads its officials to optimally choose actions that average to ψ_j. Further, this average behavior will be stable against small perturbations in perceptions or actions if, for ϕ close to ϕ_j, the averaged optimal response to ϕ leads perceptions to be revised in the direction of ϕ_j. Formally,

[5]In the body of the paper we assume that all officers perceive exactly the average behavior in their department and in the wider society, and compute their perceptions with a bias towards the departmental average. In an appendix, we discuss how the results are modified if we assume a more realistic process where each officer forms his perception based on a number of observations of actions by other agents, with a bias towards his departmental colleagues.

Definition 1 (Organizational focus) *For a given θ we define an organization's focus as a ψ_j with an associated ϕ_j such that*

$$(\psi_j, \phi_j) = (\Psi_j(\phi_j, \theta),\ \Phi(\psi_j, \theta)) \tag{4}$$

Further, a focus is stable if there is a neighborhood $N(\phi_j)$ of ϕ_j such that, for $\tilde{\phi} \in N(\phi_j) \bigcap [0,1]$, $\Psi(\tilde{\phi}, \theta) \in (\tilde{\phi}, \phi_j]$.

In other words, at a stable organizational focus, ψ_j and ϕ_j are consistent with each other via the optimization behavior of the officials, and a small perturbation in actions or perceptions from (ψ_j, ϕ_j) is self-correcting.[6]

An *equilibrium* for the economy is a set of stable organizational foci $\psi_j,\ j = 1, ..., J$ and a social norm θ, such that ψ_j is a focus for j given θ, and θ is the economy-wide average level of corruption that results from the foci $\psi_j, j = 1, ..., J$. We assume

$$\theta = \frac{\sum_j [N_j \psi_j]}{\sum_j N_j}, \tag{5}$$

i.e., it is the average of the departmental corruption levels weighted by their population weights. Thus an equilibrium is a set of organizational foci that are mutually consistent.

At an organizational focus, by (2) all officials within department j have the same perception ϕ_j, since ϕ_j is determined by the economy-wide variable θ and the department-level variable ψ_j. Thus the moral stances of officials within a department will cluster around the common mean ψ_j, deviating from the mean according to the extent to which a particular official faces corruption opportunities greater or less than the average.

To simplify the exposition and focus on the essential intuition of the model, *we will henceforth assume that all officials in all departments face the same corruption opportunities*, i.e.,

$$b_{ij} = b \quad \forall i, j.$$

It follows that, within a given department j, all officials will choose the same level of corruption,

$$\psi_{ij} \ = \ \Psi(\phi_j) = \arg\max_{\psi_{ij}} U(\psi_{ij} b) - C(\psi_{ij}, \phi_j) \tag{6}$$

which is also the average level of corruption in the department, i.e.,

$$\psi_j \ \equiv \ \Psi(\phi_j) \tag{7}$$

Allowing the b_{ij}s to vary across i and j would yield greater variations in corruption within and between departments, enhancing our results. An examination of (3) and (6) will show that nothing further is lost from this simplification.

[6]A focus is a dynamically stable partial equilibrium for the department. To avoid confusion we reserve the term "equilibrium" to refer to a general equilibrium for the economy.

2.2. Intuition

For given θ the focus of a department j is determined by the three equations, (2), (6) and (7). Note that (6) is in fact a set of N_j identical equations determining the corruption levels of the individual officials. By virtue of (7), this further reduces to two equations in the two variables ψ_j and ϕ_j.

From (2) it follows that $\Phi(\psi_j, \theta)$ lies between ψ_j and θ, has a positive slope less than unity with respect to ψ_j, and intersects the diagonal of the unit square at $\psi_j = \theta$.

The variable ψ_j is constrained to lie in the unit interval. At an interior optimum, the individual official's choice ψ_{ij} satisfies:

$$bU'(\psi_{ij}b) = C'(\psi_{ij}) \tag{8}$$

Extremal choices must satisfy corresponding inequality conditions. By (7) it follows that ψ_j satisfies the same condition.

When the solution to (6) is interior, (8) is satisfied with equality and we can derive the slope of ψ_j with respect to ϕ as:

$$\frac{\partial \Psi(\phi)}{\partial \phi} \;=\; \frac{C_{\psi\phi}}{b^2 U'' \;-\; C_{\psi\psi}} \tag{9}$$

Given the assumptions on the partials of the cost function and the utility function, it follows that $\Psi(\phi)$ is positively sloped. To further pin down the shape of the function $\Psi(\phi)$ requires assumptions about the third partials and cross-partials of the utility and cost functions. We will eschew this; instead in the closed-form model of the next section we show that there exist reasonable conditions under which it has a shape that yields satisfactory results.

Organizational foci

Figure 2.2 draws two representations of the graphs of $\Psi(\phi)$ and $\Phi(\psi, \theta)$, with ϕ on the horizontal axis.

If (ϕ_j, ψ_j) lies above the graph of $\Psi(\phi)$, then the individual choices leading to ψ_j are too high for optimum given the perception. This leads officials to reduce their corruption choices, hence (ϕ_j, ψ_j) must adjust vertically downwards. conversely if (ϕ_j, ψ_j) lies below the curve, individual actions will adjust ψ_j upwards.

We draw $\Phi(\psi_j, \theta)$ it with ϕ on the horizontal axis so it has a positive slope steeper than unity. Note that if (ϕ_j, ψ_j) is to the left of this curve, then the perception ϕ_j is too low given ψ_j and θ, so the adjustment in perception moves (ϕ_j, ψ_j) horizontally to the right. The reverse is true if (ϕ_j, ψ_j) is to the right of the curve.

A configuration (ψ, ϕ) is a potential organizational focus if the graphs of $\Psi(.)$ and $\Phi(., \theta)$ intersect at (ψ, ϕ), or if (8) is satisfied with inequalities corresponding to corner solutions. However, only some of these solutions are stable (se Definition 1), as discussed below, and hence candidates for continuing organizational cultures.

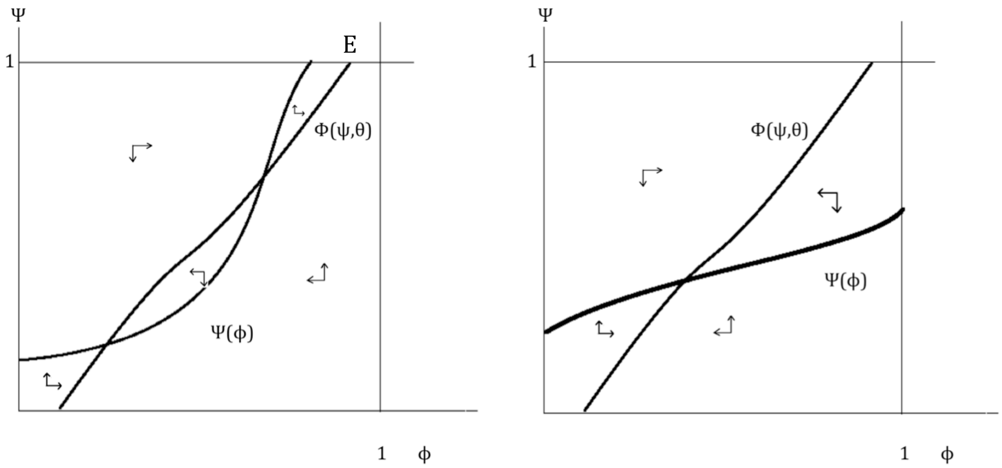

Figure 1: Perceptions and behavior, two representations

Interior solutions: Given the adjustment dynamics discussed above, it follows immediately that if the two curves intersect at an interior point (ψ^*, ϕ^*) where $\Psi(.,.)$ is less steep than $\Phi(.,.)$, then (ψ^*, ϕ^*) is a stable equilibrium of the system, and hence is a possible departmental focus. On the other hand if $\Psi(.,.)$ is steeper than $\Phi(.,.)$, then (ψ^*, ϕ^*) is not stable, an infinitesimal deviation in either individual actions or perceptions will lead to a movement away from this point. Hence such an intersection does not constitute a focus.

Corner solutions: If $\Psi(.)$ lies entirely above $\Phi(.,.)$, or if the rightmost intersection of the two curves is unstable, then $(\phi, \psi) \;=\; (\Phi(1, \theta),\; 1)$ is a stable focus. This is point E in the left panel of figure 2.2. Under the given conditions the graph of $\Psi(.)$ intersects the horizontal line at $\psi = 1$ to the left of the graph of $\Phi(.,.)$. Thus $\psi = 1$ is a corner solution to (6) when $\phi = \; \Phi(1, \theta)$. Clearly, (2) is satisfied at E, given $\psi = 1$. To see that E is stable, note that point D, the rightmost intersection between the two curves, is unstable. Thus a small displacement from E that remains above and to the right of D sets up dynamics that converge back to E.

Finally, note that $\Psi(.) \geq 0$ at $\phi = 0$, and $\Phi(.,.) \geq 0$ at $\psi = 0$. Thus at $\phi = 0$ the graph of $\Psi(.)$ lies (weakly) above the graph of $\Phi(.,.)$.

Together these observations imply that there are three possible configurations of foci:

1. The graph of $\Psi(.,.)$ remains above that of $\Phi(.,.)$ for all relevant values of ϕ, in which case $(\phi, \psi) \;=\; (\Phi(1, \theta),\; 1)$ is the only stable organizational focus.
2. The graph of $\Psi(.,.)$ cuts that of $\Phi(.,.)$ from above only once, and then remains below it, in which case the unique intersection is the sole organizational focus.

If there is a unique focus, then it is sustainable in general equilibrium only when at the intersection $\psi = \phi = \theta$. Note that all organizations would then have this same focus, in turn justifying $\theta = \psi$. If the unique intersection is interior, then all organizations choose the same ψ_j, and hence we must again have $\psi_j = \phi_j = \theta$ for all j. To determine the full equilibrium for the economy requires us to endogenize θ, which is addressed in Section 4.

The unique focus may be at $(0, 0)$ if $\theta = 0$. This is an equilibrium with no corruption in the economy. Alternatively the unique focus may be at $(\Phi(1, \theta),\ 1)$. Then all organizations must choose $\psi = 1$, implying $\theta = 1$, so the economy is completely corrupt. In turn this implies $\phi_j = 1$ for all j.

3. The graphs may intersect more than once in the interior. The leftmost intersection must produce a stable focus (since at $\phi = 0$ the graph of $\Psi(.)$ lies (weakly) above the graph of $\Phi(.,.)$). The next intersection must therefore be unstable (the graphs intersect in the "wrong" direction), and every odd-numbered intersection must be stable. Finally, if the rightmost intersection is unstable then there is an extreme focus at $(\Phi(1, \theta),\ 1)$.

When there are multiple foci, θ is a weighted average of the corruption levels chosen by organizations that locate at the various foci. It is important to emphasize that the multiple foci still constitute a single equilibrium for the economy, in which different organizations locate at different levels of corruption.

Organizational change and economywide dynamics

If some organizations were to relocate from one focus to another for any reason, this would result in a change in θ. In turn this shifts the graph of $\Phi(., \theta)$, throwing the remaining organizations out of focus. The resulting dynamics could possibly produce significant changes in the profile of corruption in the economy.

As an illustration, consider Figure 2.2. Suppose the economy is in equilibrium with $\theta = \theta_0$, with corresponding masses of organizations located at the stable foci A, C and E. Now, consider a large organization j located at E. Let the perception of corruption ϕ_j at this organization undergo a downward perturbation.[7] If the perturbation is large enough, the employees of j may initially find themselves out of focus to the left of point D, and in the ensuing adjustment converge to C. But in turn this lowers θ from θ_0 to θ_1, and hence shifts the graph of $\Phi(., \theta)$ to the left (dotted line). This results in organizations at A and C falling out of focus, leading them to converge to A' and C'. In turn this further lowers θ to θ_2 and causes a further shift in the graph of $\Phi(., .)$. If the graph shifts sufficiently (i.e., if the initial relocations were made by sufficiently large or prominent organizations), the final outcome may be as in the right-hand panel of the figure, where the only possible focus is A''.

We can now summarize the important insights of the present model.

[7] Such a change in perception may occur, for instance, if another organization with a very low culture of corruption physically relocates near organization j. A change in perception may also be engineered by a change in governance, as in the examples of Hong Kong and Phillippines described by Klitgaard and cited in the introduction.

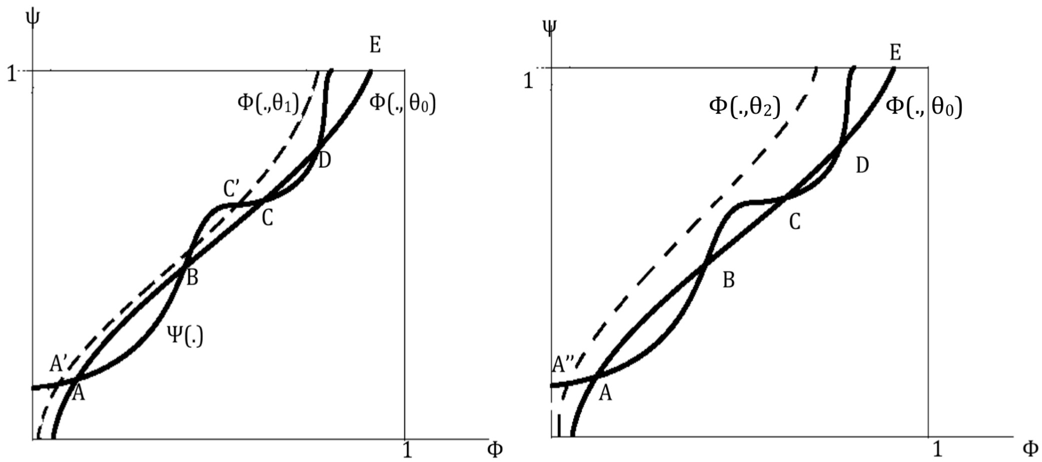

Figure 2: Organizational changes and effects on the economy

(a) Individual organizations with diverse levels of corruption can co-exist in equilibrium in an economy.
(b) Internal changes in individual organizations can shift them from high-corruption to low-corruption foci.
(c) Individual changes can set in motion cascades that result in large and precipitous changes in the economy, as in coordination models.

2.3. Closed-form model

In the remainder of the paper, we establish the general results within a closed-form model. This makes the exposition simpler, and allows us to explicate the results in terms of variables (α and β, defined later) that have intuitive interpretations. The explicit model is:

$$U(\psi_{ij}b) = \psi_{ij}b \tag{10}$$

$$C(\psi_{ij}, \phi_j) = \frac{1}{2}z(\frac{\psi_{ij}}{\phi_j})^2 - \frac{1}{2}z \tag{11}$$

$$\Phi(\psi_j, \theta) = \alpha\psi_j + (1-\alpha)\theta \tag{12}$$

Where α is a measure of the social isolation of individuals—the extent to which their perceptions of norms are colored by their immediate peers—and z is a parameter denoting the intensity of cost felt from behavior that deviates from the perceived norm.

Note that the moral cost C is strictly convex and increasing everywhere in its range. In our formulation it is zero when $\psi_{ij} = \phi_j$, i.e., when the official behavior is in line with his perception of the social norm, and becomes negative when he acts “more morally” than the norm. This is in keeping with our interpretation of the cost, though not necessary for the results. Further, note that given a strictly convex and increasing cost function, a linear utility function does not compromise generality, it is the concavilty of utility net of cost that drives the results.

3. Corruption at the department level

In this section we establish the possible organizational foci assuming an exogenous economy-wide average corruption level $\theta \in [0, 1]$. The only property of θ that we use in this section is that it is a weighted average of possible organizational focus levels of ψ, and hence must be a convex combination of these foci.

We show that the number and positions of possible organizational foci depends on the relative value of the parameters α, β and θ (Observation 1). At most, one focus (ψ^l) is interior and stable (Lemma 1) where the other interior solution (ψ^h) acts as an unstable tipping point between total corruption and a low levels of corruption (Corollary 2). The corner solution of complete corruption (ψ^H) is also a possible (Lemma 3) or unique (Lemma 4) stable focal point under certain conditions. These are elucidated more formally below. Proofs are in the Appendix.

If the (common) optimization problem of the officials in department j has an interior solution then ψ_j must satisfy

$$\psi_j = \frac{b}{z}\phi_j^2 \tag{13}$$

If this yields a value for ψ_j outside the unit interval, then the corresponding inequality conditions will apply. Note that different organizations may nevertheless settle at different stances.

Define $\beta = \frac{b}{z}$. Given the interpretations of b and z, this exercise is relevant only if $\beta > 0$. Accordingly we assume this is true.

Henceforth we suppress the subscript j when it does not cause confusion. Substituting (13) in (12) and rearranging gives

$$\alpha\beta\phi^2 - \phi + (1-\alpha)\theta = 0 \tag{14}$$

The equation (14) has two roots for ϕ, given by

$$\phi^h = \frac{1 + \sqrt{1 - 4\alpha(1-\alpha)\beta\theta}}{2\alpha\beta} \tag{15}$$

$$\phi^l = \frac{1 - \sqrt{1 - 4\alpha(1-\alpha)\beta\theta}}{2\alpha\beta} \tag{16}$$

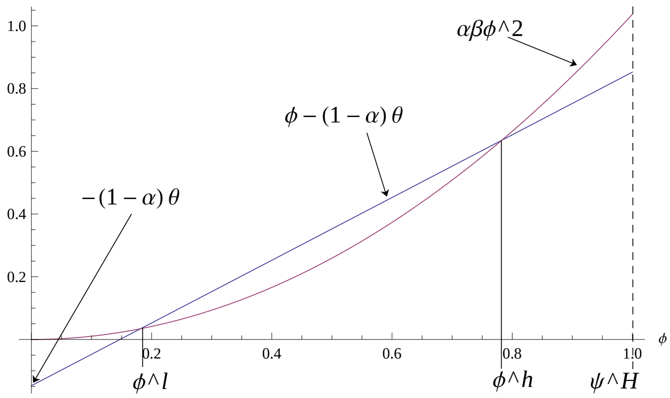

Figure 3: Solutions for equation (14)

These are the two possible values for the perceived norm for any department at an organizational focus. The solutions are illustrated in Figure 3, as the intersections between the line $y = \phi - (1-\alpha)\theta$ (approximately denoting the costs to corruption), and the curve $y = \alpha\beta\phi^2$ (approximately denoting the benefits to corruption). The latter curve is anchored at zero and pivots upwards anticlockwise as β increases, raising ϕ^l and lowering ϕ^h. If β is sufficiently large then the curve lies entirely above the line, and (14) has no real solutions—any given department can only focus on complete corruption.

Observation 1 *If individuals' perception of society is imperfect (α is interior) and the general social norm is below a particular level ($\theta < \frac{1}{4\alpha(1-\alpha)}\frac{1}{\beta}$), then ϕ^h and ϕ^l are real and distinct, and $\phi^h > \phi^l$. If α is zero (everyone observes society perfectly) or unity (individuals cannot see beyond their immediate peers) then (14) has a single solution equal to θ or $\frac{1}{\beta}$, respectively.*

First consider the case where the roots are real and distinct. From (13) and (14), these correspond to distinct levels of average corruption:

$$\psi^h = \frac{1}{\alpha}[\phi^h - (1-\alpha)\theta]$$
$$\psi^l = \frac{1}{\alpha}[\phi^l - (1-\alpha)\theta] \tag{17}$$

We show below that a third candidate for an organizational focus is the stance $\psi = 1$, which by (13) and the corresponding inequality condition requires $\phi \geq \frac{1}{\sqrt{\beta}}$. However, the perception ϕ itself is given by a weighted average of θ and ψ, as in (12). For convenience define:

$$\psi^H = 1, \quad \phi^H = \alpha + (1-\alpha)\theta \tag{18}$$

Note that ϕ^H cannot exceed unity.

Given θ, a solution $\psi^k,\ (k = l, h, H)$ with the corresponding perception ϕ^k is a candidate for an organizational focus if both ϕ^k and ψ^k lie in the unit interval.

We can now identify the potential organizational foci as defined in Definition 1. Given an average economy-wide corruption level θ, an average corruption level $\psi_j \in [0, 1]$ in organization j results in a perception $\phi_j = \alpha\psi_j + (1-\alpha)\theta$, according to (12). Clearly $\phi_j \geq 0$. By (13) and the boundary constraints, the optimal ψ_j corresponding to ϕ_j must in turn satisfy

$$\psi_j = \begin{cases} \beta\phi_j^2 & if \quad \beta\phi_j^2 \in [0,1] \\ 1 & if \quad \beta\phi_j^2 > 1 \end{cases}$$

A ψ_j with an associated ϕ_j that satisfies the above conditions is a stable focus if it is stable in the sense of Definition 1.

First we address the question of stability. Lemma 1 below shows that, if (14) has real and distinct solutions, then the lower solution ϕ^l corresponds to a stable focus, but the higher one ϕ^h does not. In other words, if members of an organization have a (disequilibrium) perception close to ϕ^l, then their responses to that perception will lead the organization to settle at a focus ϕ^l. However, if their perception is iclose to (but not equal to) ϕ^h, then their collective responses will lead the organization further away from ϕ^h. This follows directly from Lemma 1 below, and is made explicit in the subsequent proposition. All proofs are in the appendix.

Lemma 1 *If both ϕ^k and ψ^k, $(k = h, l)$, lie in the unit interval then ψ^l is a stable focus, but ψ^h is not stable.*

Proposition 2 (Corollary to Lemma 1) *Let* (14) *have real solutions ϕ^l and ϕ^h, with $\phi^l \leq \phi^h$. Suppose an organization currently has a perception $\tilde{\phi} \in [0, 1)$ with a corresponding stance $\psi = \beta(\tilde{\phi})^2$. If $\tilde{\phi} < \phi^l$ then the stance induces an increase in the perception, if $\tilde{\phi} \in (\phi^l, \phi^h)$ then it induces a decrease in the perception, and if $\tilde{\phi} > \phi^h$ it induces an increase in the perception.*

Note that the formal statement of Proposition 2 does not require that the solutions lie in the unit interval. It also does not presuppose that $\tilde{\phi}$ is an equilibrium perception.

Since $\psi^h = \beta(\phi^h)^2$ is unstable, we will disregard it as a candidate for an organizational focus. If $\psi^l < 1$ and an organization locates (out of equilibrium) at a corruption stance $\tilde{\psi} < \psi^h$, this gives rise to a perception $\tilde{\phi} < \phi^h$, and by Corollary 2 officials in the organization must revise their perception and stance downwards until it reaches ψ^l. Similarly, if $\tilde{\psi} > \psi^h$ then the organization must revise upward towards unity, a value of $\psi \in (\psi^h, 1)$ cannot be an organizational focus. Once the organization reaches a perception $\psi^H = 1$, it cannot choose a higher stance. However, ψ^H can be an organizational focus if θ is sufficiently large, or if ϕ^l is large, as Lemma 3 shows.

Lemma 3 *(a) If $\phi^l < \phi^H$ then ψ^H is a stable organizational focus if and only if $\beta > \max\{1, \frac{1}{4\alpha^2}\}$ and $\theta \geq \frac{1-\alpha\sqrt{\beta}}{(1-\alpha)\sqrt{\beta}}$.*
(b) If $\phi^l \geq \phi^H$ then ψ^H is a stable organizational focus.

ϕ^H and ψ^H were defined in (18) earlier. Extending the argument for part (b) above, it is easy to see that if β is sufficiently large so that in Figure 3 the curve $y = \alpha\beta\phi^2$ lies entirely above the line $y = \phi - (1-\alpha)\theta$, then ϕ^H is stable against downward deviations, and ψ^H is the only possible organizational focus. We summarize this and omit the proof.

Lemma 4 *if (14) does not have real solutions, then ψ^H is the unique stable organizational focus.*

The preceding discussion suggests that there are at most two candidates for stable organizational foci, ψ^l and ψ^H. From Lemma 4 it follows that if ψ^l is not an organizational focus, then ψ^H is a focus, hence at least one focus must exist.

If both foci exist, then organizations in the economy will locate at one or the other, and the average corruption level θ in the economy will be a weighted average of the two. We will state this formally in the next section. For the moment note that since $\psi^H = 1$ we have $\theta \leq \psi^H$, so if ψ^l is also a focus then it must satisfy $\psi^l \leq \theta$.

4. A general equilibrium for the economy

The analysis in the previous section leads to the conclusion that there is always at least one focus on which organizations can coordinate. Depending on the parameters either ψ^l or ψ^H or both may be potential foci. We will further see below that, for the same parameters (α and β), the economy may have multiple potential sets of foci.[8]

The equilibrium for the economy is defined by the set of organizational foci, the distribution of organizations between them, and the resulting economy-wide average level of corruption θ. If there is only one organizational focus ψ^* then all organizations will locate there, and θ must also equal ψ^*. If both ψ^l and ψ^H are organizational foci, then individual organizations will locate at one or the other, and θ will be a weighted average of the two. This leads us to the definition of the economy-wide equilibrium.

Let $\pi = (\pi^l, \pi^H)$ be the set of employment-weighted fractions of organizations that locate at ψ^l and ψ^H respectively, so that $\pi^k \geq 0$ *for* $k = l, H$ and $\pi^l + \pi^H = 1$.

Definition 2 (Equilibrium) *An* equilibrium *for the economy consists of an aggregate corruption level θ, an associated nonempty set of foci $\Psi \subseteq \{\psi^l, \psi^H\}$, and a distribution π that satisfy:*

a. $\pi^k > 0$, $k = l, H$ *only if ψ^k is an organizational focus given θ.*

b. $\theta = \sum_{k:\psi^k \in \Psi} \pi^k \psi^k$.

Depending on the values of the parameters α and β, an economy may have three different kinds of equilibria. Figure 4 shows the regions of the parameter space at which each kind of

[8]The functional forms chosen produce a saddle shaped map of potential focal points, the shape of which changes as θ varies.

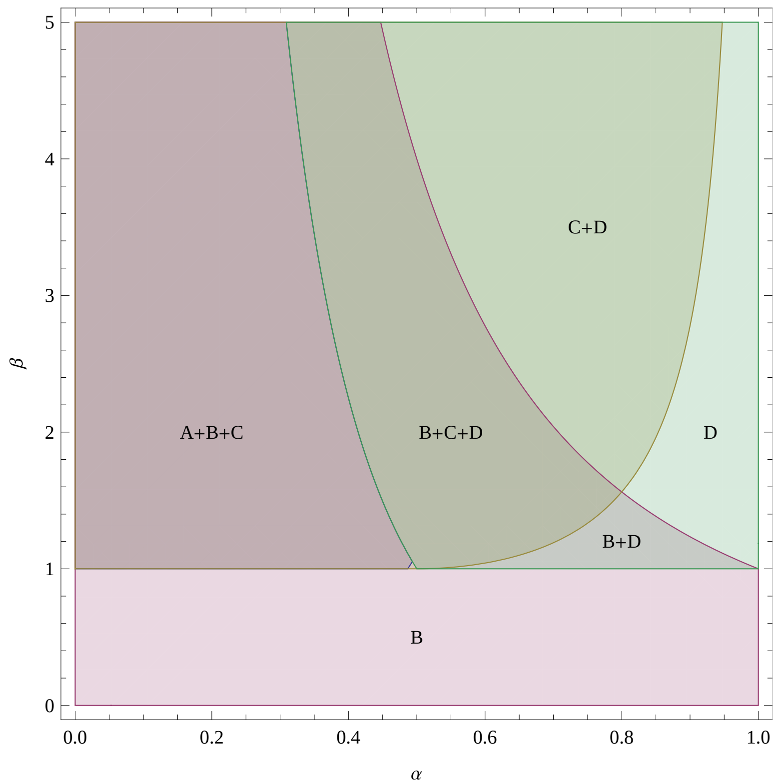

Figure 4: Equilibria at different parameter values

equilibrium is possible. In regions A and B, there is an equilibrium in which all organizations in the economy coordinate on an organizational focus at ψ^l. In region C, there is an equilibrium in which all coordinate on a focus at ψ^H. In region D there is an equilibrium that incorporates both foci, with some organizations locating at ψ^l while others locate at ψ^H. These regions overlap, indicating that there are ranges in the parameter space where multiple equilibria exist. We now investigate conditions under which each type of equilibrium obtains.

First note that, if there is only one organizational focus $\psi^k, k = l$ or H, then θ must equal ψ^k. This leads to the following two propositions.

Proposition 5 *(i) If in an equilibrium ψ^l is the sole organizational focus, then either $\theta = \frac{1}{\beta}$ or $\theta = 0$.*
(ii) There is an equilibrium such that $\theta = \psi^l = \frac{1}{\beta}$ and ψ^l is the only organizational focus if $\alpha < \dfrac{1}{2}$ and $\beta \in [1, (\frac{1-\alpha}{\alpha})^2])$ (region A in Figure 4).
(iii) There is an equilibrium with $\theta = \psi^l = 0$, and ψ^l is the only organizational focus if $\beta < \frac{1}{\alpha^2}$ (region B in Figure 4).

Proposition 6 *There is an equilibrium such that $\theta = \psi^H = 1$ and ψ^H is the only organizational focus if either (a) $\beta > \frac{1}{4\alpha(1-\alpha)}$ or (b) $\alpha < \dfrac{1}{2}$ and $\beta \in (1, \frac{1}{2\alpha}]$ (region C in Figure 4).*

Finally, we turn to the conditions under which there is an equilibrium with two distinct organizational foci, ψ^l and ψ^H, both with positive weights.

Proposition 7 (Multiple equilibria) *If (14) has real roots, then there is an equilibrium such that both ψ^l and ψ^H are organizational foci and a positive fraction of organizations locate at each focus if either*
(i) $\alpha < \dfrac{1}{2}$ and $\beta > [\frac{1-\alpha}{\alpha}]^2$, or
(ii) $\alpha \geq \dfrac{1}{2}$, $\beta > 1$.
(region D in Figure 4.)

In summary, an equilibrium is the set of the stable organizational foci, the proportion of departments at each focus, and the resulting average corruption level. At given parameter values, an equilibrium is comprised of either a unique interior focal point for all organizations (Proposition 6), a unique focal point at complete corruption for all organizations (Proposition 7) or both of these with a proportion of organizations focusing on each (Proposition 8).

Multiple equilibria can exist for a wide range of parameter values. For example in figure 4, if α and β are in the region marked B+C+D, there is an equilibrium in which all organizations settle at a focus with $\psi = 0$ (Proposition 5), an equilibrium in which all organizations focus on $\psi = \psi^H = 1$ (Proposition 6), and an equilibrium in which some organizations settle at an interior level ψ^l while others settle at ψ^H (Proposition 7).

In particular, in Proposition 7, we have our description of an equilibrium in which otherwise identical departments in the same economy focus on very different patterns of corrupt behavior. The *possibility* of this state is determined by prevailing incentives and the degree of social isolation, and its *existence* is determined by officials' perceptions of the rest of society. In the state where it does exist, the level of corrupt behavior on which a particular organization focuses will be determined by officials' perceptions of their colleagues' beliefs, as discussed in Section 3. Hence in this equilibrium we can see how distinct organizational cultures might establish and persist over time in an economy with uniform incentives and homogeneous agents.

5. Conclusion and policy implications

Whereas most economic corruption models focus on structural incentives, our model investigates how peer effects can produce different but persistent cultures of corruption. We examine officials' decisions when they must contend with their own personal morals and also their perceptions of what is normal. We find multiple organizational foci can exist. Furthermore it is feasible for these foci to persist over time in the economy as officials' perceptions and decisions create organizational cultures that propagate the behavior.

It should be emphasized that we used a simple model, and hence derived correspondingly stark results. However, many of the simplifications generalize readily. In particular, we draw attention to one consequence of our simplifications; in any equilibrium, each official in each of the organizations located at a given focus is corrupt to exactly the same degree. As a result, all of these organizations are equally corrupt. Clearly this will not be the case if different officials in the same department have (i) different levels of exposure to the broader economy (i.e., different αs) and (ii) different moral costs z for deviating from what they consider the true moral line, and (iii) different opportunities for corruption (different b_{ij}s). If this is the case, then even within a department each official will avail his corruption opportunities to a different degree, and indeed the departments located near a given focus will in fact display different average levels of corruption. Further, the actual perceived levels of corruption in different departments will vary according to the vantage point of the observer (whether she dealt with a more or less corrupt official). Admitting such variations does not compromise the validity of our model.

It is of greater importance to note that our analysis provides very definite and sometimes novel perspectives on policy. Consider an economy in an equilibrium with high levels of corruption, either because total corruption is the only organizational focus or because there are multiple foci and a high proportion of organizations have settled on this option. Our results suggest four linked implications that could inform anti-corruption policy in this economy:

1. In accordance with preceding literature, reducing the relative benefits of corruption (i.e., reducing β by decreasing the rewards of corruption b or increasing its cost z) will reduce corruption. The mechanism at work here is that as the benefits are reduced, the possible equilibria change: in extreme cases (where the benefits to corruption are all but eliminated) the possibility of a highly corrupt focal point in equilibrium is completely removed. In other cases, reducing the benefits to corruption may introduce a tipping point ($\phi^h < \phi^H$) below which organizations will focus on lower levels of corruption. As reductions continue further, this tipping point will increase to higher levels and thus more corrupt organizations will fall below it. In either case, aggregate corruption will decrease.

2. In the equilibrium with two focal points, significant changes in departmental corruption levels may be brought about with no change to the incentive structure but rather by targeting the social-psychological mechanisms discussed earlier. That is, significant improvements within organizations may be achieved by policies that aim to disrupt the corrupt subcultures. For example, targeting individuals' perceptions of what the norm is by discrediting prevalent rationalizations for the behavior will increase the moral costs individuals bear and could effect an organizational switch from the higher

to lower focal points. These policies will be more effective if the tipping point is higher, because they only need to push bureaucrats' expectations of others' behavior across that line, rather than drive them down all the way to the lower equilibrium.

3. This gives rise to the key observation that the battle to reduce corruption in an economy overall can be won by whittling away at the cultures of individual departments one at a time, rather than with a big push against all departments at once. In fact, by targeting policies such as those discussed above at particular (and particularly visible) organizations, policymakers may effect a change in the proportion of organizations focusing on the higher focal point in the equilibrium with two foci. At the margin, the structure of the equilibrium changes, so that complete corruption is no longer a sustainable focus for any department, regardless of whether they have been targeted by reform programs.

4. Finally, note that in Proposition 7, organizations may congregate at the high corruption equilibrium even when incentives are relatively weak ($\beta > 1$) if their degree of social isolation is high ($\alpha > \frac{1}{2}$). Here departmental corruption is driven largely by internal culture, not the state of corruption in the broader economy. Providing officials with a truer perception of the prevailing social norm will here be effective in shifting the department away from its prevailing focus. However, this is a delicate instrument, since in the departments located at the low corruption focus the same policy would have precisely the opposite effect.

In conclusion, we draw attention to our interpretation of the assumption regarding the cost of corruption. Our interpretation is that the decision to be corrupt is not entirely a function of narrowly economic costs and benefits, it is also in large part driven by moral factors that derive their force from social norms, which in turn are formed by common and accepted practice. We believe that incorporating this aspect of the decision in economic modelling of the phenomenon of corruption is potentially beneficial for policy formulation.

References

Akerlof, George (1976), 'The economics of caste and of the rat race and other woeful tales', *The Quarterly Journal of Economics* **90**(4), 599–617.

Akerlof, George A. and William T. Dickens (1982), 'The economic consequences of cognitive dissonance', *The American Economic Review* **72**(3), 307–319.

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

Arendt, Hannah (1977), *Eichmann in Jerusalem: a report on the banality of evil*, Penguin Books, London.

Ashforth, Blake E. and Vikas Anand (2003), 'The normalization of corruption in organizations', *Research in Organizational Behavior* **25**, 1–52.

Bardhan, Pranab (1997), 'Corruption and development: A review of issues', *Journal of Economic Literature* **35**(3), 1320–1346.

Basu, Kaushik (2000), *Prelude to Political Economy: A Study of the Political Foundations of Economics*, Oxford: Oxford University Press.

Blackburn, Keith, Niloy Bose and M. Emranul Haque (2006), 'The incidence and persistence of corruption in economic development', *Journal of Economic Dynamics & Control* **30**, 2447–2467.

Bose, Gautam and Shubhashis Gangopadhyay (2009), 'Intermediation in corruption markets', *Indian Growth and Development Review* **2**(1), 39–55.

Coleman, James S. (1990), *Foundations of Social Theory*, Harvard University Press, Cambridge, MA.

Cressey, Donald R. (1971), *Other People's Money: A study in the social psychology of embezzlement*, Wadsworth, Belmont, CA.

Elster, Jon (1989), 'Social norms and economic theory', *The Journal of Economic Perspectives* **3**(4), 99–117.

Fischer, Paul and Steven Huddart (2008), 'Optimal contracting with endogenous social norms', *American Economic Review* **98**(4), 1459–75.

Glaeser, Edward L, Bruce Sacerdote and José A Scheinkman (1996), 'Crime and social interactions.', *Quarterly Journal of Economics* **111**(2).

Ichino, Andrea and Giovanni Maggi (2000), 'Work environment and individual background: explaining regional shirking differentials in a large italian firm', *The Quarterly Journal of Economics* **115**(3), 1057–1090.

Klitgaard, Robert E. (1988), *Controlling Corruption*, University of California Press, Berkeley and Los Angeles, CA.

Laub, John H. (2006), 'Edwin H. Sutherland and the Michael-Adler Report: Searching for the soul of criminology seventy years later', *Criminology* **44**(2), 235–257.

Murphy, Kevin M., Andrei Shleifer and Robert W. Vishny (1993), 'Why is rent-seeking so costly to growth?', *The American Economic Review* **83**(2), 409–414.

Nabin, Munirul Haque and Gautam Bose (2008), 'Partners in crime: Collusive corruption and search', *The B.E. Journal of Economic Analysis & Policy* **8**(1).

Rabin, Matthew (1994), 'Cognitive dissonance and social change', *Journal of Economic Behavior & Organization* **23**, 177–194.

Rasmusen, Eric (1996), 'Stigma and self-fulfilling expectations of criminality', *Journal of Law and Economics* pp. 519–543.

Sardan, J. P. Olivier de (1999), 'A moral economy of corruption in Africa?', *The Journal of Modern African Studies* **37**(1), 25–52.

Schneider, Patrick (2011), Everyone was doing it: Culture, Rationalisation and Corruption, Honours thesis, School of Economics, University of New South Wales, Sydney.

Wade, Robert (1982), 'The system of administrative and political corruption: Canal irrigation in South India', *Journal of Development Studies* pp. 287–328.

Wade, Robert (1985), 'The market for public office: why the indian state is not better at development', *World Development* **13**(4), 467–497.

Appendix A Proofs

Proof of Lemma 1

Suppose that both ϕ^l and ϕ^h lie in the unit interval. Consider a perceived corruption level $\tilde{\phi} = \phi^k + \Delta, \ \ \Delta \neq 0$ that deviates by a small amount Δ from one of the solutions. Given θ, this leads to a choice of corruption level $\tilde{\psi} = \beta\tilde{\phi}^2 = \beta[(\phi^k)^2 + 2\Delta\phi^k + \Delta^2]$. In conjunction with θ, $\tilde{\psi}$ must in turn give rise to a new perception $\phi' = \alpha\beta[(\phi^k)^2 + 2\Delta\phi^k + \Delta^2] + (1-\alpha)\theta$, which reduces to $\phi' = \alpha\beta[2\Delta\phi^k + \Delta^2] + \phi^k$ using (14). The perception ϕ^k is stable if

$$\begin{aligned} |\phi' - \phi^k| &< |\tilde{\phi} - \phi^k| \\ \Rightarrow |2\alpha\beta\Delta\phi^k + \alpha\beta\Delta^2| &< |\Delta| \\ \Rightarrow |2\alpha\beta\phi^k + \alpha\beta\Delta| &< 1 \end{aligned} \tag{19}$$

which in the limit as $\Delta \to 0$ requires $2\alpha\beta\phi^k < 1$. It follows immediately from (15) and (16) that the smaller solution ϕ^l is stable while the larger, ϕ^h, is not. ■

Proof of Lemma 3

(a) First, by (13), for ψ^H to be an optimal stance the corresponding perception must satisfy $\beta(\phi^H)^2 > 1$ which implies $\phi^H \geq \frac{1}{\sqrt{\beta}}$. However, according to (18) $\phi^H = \tilde{\phi} = \alpha + (1-\alpha)\theta$. This exceeds $\frac{1}{\sqrt{\beta}}$ if and only if $\theta \geq \frac{1-\alpha\sqrt{\beta}}{(1-\alpha)\sqrt{\beta}}$.

Next, the perception ϕ^H is stable against upward perturbations because ψ cannot exceed unity. Since $\phi^l < \phi^H$, by Corollary 2 ψ^H is stable against downward deviations only if $\phi^h < \phi^H$. By the earlier part of the proof, a sufficient condition for this is $\phi^h < \frac{1}{\sqrt{\beta}}$.

From (15), $\phi^h < \frac{1}{\sqrt{\beta}} \Leftrightarrow \sqrt{1-4\alpha(1-\alpha)\beta\theta} < 2\alpha\sqrt{\beta}-1$. Since the LHS is positive (real roots) the RHS must be positive, which requires $\beta > \frac{1}{4\alpha^2}$. Squaring both sides then again yields $\theta \geq \frac{1-\alpha\sqrt{\beta}}{(1-\alpha)\sqrt{\beta}}$, hence the necessary condition for $\psi^H = 1$ is sufficient for $\phi^h < \phi^H$.

Finally, note that $\theta \geq \frac{1-\alpha\sqrt{\beta}}{(1-\alpha)\sqrt{\beta}}$ requires that $\frac{1-\alpha\sqrt{\beta}}{(1-\alpha)\sqrt{\beta}} < 1$, for which either $\beta > 1$ or the numerator is non-positive implying $\beta > \frac{1}{\alpha^2} > 1$. Thus β must in fact exceed $\max\{1, \frac{1}{4\alpha^2}\}$.

(b) Follows from Corollary 2. ■

Lemma 8 *When (14) has real roots, $\psi^l \leq \theta$ if either (i) $\alpha \geq \dfrac{1}{2}$ or (ii) $\theta \leq \frac{1}{\beta}$.*

Proof of Lemma 8

From (12) it follows that $\psi^l \leq \theta$ if and only if $\phi^l \leq \theta$. From (16), this requires

$$\sqrt{1-4\alpha(1-\alpha)\beta\theta} \geq 1-2\alpha\beta\theta$$

It follows that:
(a) If $\theta < \frac{1}{2\alpha\beta}$ then the RHS is positive, squaring both sides and simplifying yields $\theta < \frac{1}{\beta}$.
(b) If $\theta \geq \frac{1}{2\alpha\beta}$ then the RHS is non-positive and the condition is satisfied.

Note also:
(c) $\theta \geq \frac{1}{2\alpha\beta}$ implies $4\alpha(1-\alpha)\beta\theta \geq 2(1-\alpha)$. Real roots require $4\alpha(1-\alpha)\beta\theta \leq 1$ so we must have $2(1-\alpha) \leq 1$ or $\alpha \geq \dfrac{1}{2}$.
(d) $\frac{1}{\beta} \lesseqqgtr \frac{1}{2\alpha\beta}$ as $\alpha \lesseqqgtr \dfrac{1}{2}$.

(i) Let $\alpha \geq \dfrac{1}{2}$. Then $\frac{1}{\beta} \geq \frac{1}{2\alpha\beta}$. So if $\theta < \frac{1}{2\alpha\beta}$ then $\theta < \frac{1}{\beta}$ and $\phi^l < \theta$ by (a), and if $\theta \geq \frac{1}{2\alpha\beta}$ and (14) has real roots then $\phi^l < \theta$ by (b).

(ii) By the first part of (i), $\theta < \frac{1}{\beta} \rightarrow \phi^l < \theta$ if $\alpha \geq \dfrac{1}{2}$. If $\alpha < \dfrac{1}{2}$, then $\frac{1}{\beta} < \frac{1}{2\alpha\beta}$ by (d), so $\theta < \frac{1}{\beta} \Rightarrow \phi^l < \theta$ by (a).

Further, if $\alpha < \dfrac{1}{2}$ and $\theta \geq \frac{1}{2\alpha\beta}$ then (14) does not have real roots by (c), so (b) does not apply. ■

Proof of Proposition 5

(i) Since ϕ^l is a convex combination of θ and ψ^l, if ψ^l is the only O-equilibrium then we must have $\theta = \psi^l = \phi^l \Rightarrow \beta(\phi^l)^2 = \phi^l$ which implies $\phi^l = 0$ *or* $\frac{1}{\beta}$.
(ii) Let $\theta = \phi^l = \psi^l = \frac{1}{\beta}$, which implies that $\beta \geq 1$. When $\theta = \frac{1}{\beta}$, the two roots of (14 are $\frac{1}{\beta}$ and $\frac{1-\alpha}{\alpha\beta}$. Since we want $\phi^l = \frac{1}{\beta}$ we need this to be the smaller root, which implies

$\alpha < \dfrac{1}{2}$. Further we need that ψ^H is not an organizational focus, by Lemma 3, a necessary and sufficient condition for this is $\theta = \frac{1}{\beta} < \frac{1-\alpha\sqrt{\beta}}{(1-\alpha)\sqrt{\beta}}$ which reduces to $\beta \leq (\frac{1-\alpha}{\alpha})^2$. [9]
(iii) If $\theta = \phi^l = \psi^l = 0$ is the only equilibrium then ψ^H is not an equilibrium, so we must have $\phi^h \geq \phi^H$. Given that $\theta = 0$, we have $\phi^H = \alpha$ and $\phi^h = \frac{1}{\alpha\beta}$. Hence the required condition is

$$\phi^h \geq \phi^H \quad \Leftrightarrow \quad \frac{1}{\alpha\beta} \geq \alpha \quad \Leftrightarrow \beta \leq \frac{1}{\alpha^2}$$

■

Proof of Proposition 6

Let $\theta = 1$. If condition (a) is satisfied then (14) has no real roots hence there are no solutions other than ψ^H, so the latter is an equilibrium. If (a) is not satisfied then we need $\phi^l \geq 1$ to ensure that ψ^H is the only equilibrium. Using (16) and $\theta = 1$, this condition is equivalent to $\sqrt{1 - 4\alpha(1-\alpha)\beta} \leq 1 - 2\alpha\beta$. Since the LHS is non-negative we must have the RHS non-negative as well which requires $\beta \leq \frac{1}{2\alpha}$. Squaring both sides then yields $\beta > 1$, so we must also have $\alpha < \dfrac{1}{2}$. ■

Proof of Proposition 7

Note (a) since θ is a convex combination of ψ^l and ψ^H, and $\psi^H = 1 > \theta$, we must have $\psi^l < \theta$. Therefore by Lemma 8 either $\alpha < \dfrac{1}{2}$ and $\theta < \frac{1}{\beta}$, or $\alpha \geq \dfrac{1}{2}$.
Further, (b) ψ^H must be a stable equilibrium, so by Lemma 3 we need $\theta > \frac{1-\alpha\sqrt{\beta}}{(1-\alpha)\sqrt{\beta}}$ and $\beta > \max\{1, \frac{1}{4\alpha^2}\}$.

(i) Let $\alpha < \dfrac{1}{2}$. We need $\frac{1-\alpha\sqrt{\beta}}{(1-\alpha)\sqrt{\beta}} < \theta < \frac{1}{\beta}$. Some manipulation gives $\frac{1}{\beta} > \frac{1-\alpha\sqrt{\beta}}{(1-\alpha)\sqrt{\beta}} \Leftrightarrow \beta > [\frac{1-\alpha}{a}]^2$. Note that if $\beta > 1$ then $\frac{1}{\beta} \in (0,1)$ and hence $\frac{1-\alpha\sqrt{\beta}}{(1-\alpha)\sqrt{\beta}} < 1$, so the interval has a non-empty intersection with the range of θ. Further, since $\alpha < \dfrac{1}{2} \Rightarrow \alpha^2 < \frac{1}{4}$ so $\max\{1, \frac{1}{4\alpha^2}\} = 1$, In turn $\frac{1-\alpha}{a}]^2 > 1$ so the condition on β in (b) is satisfied. Now pick $\theta \in (\frac{1-\alpha\sqrt{\beta}}{(1-\alpha)\sqrt{\beta}}, \frac{1}{\beta})$, and define $\pi_\theta = \frac{1-\theta}{1-\psi^l}$ which implies $\theta = \pi_\theta\psi^l + (1-\pi_\theta)\psi^H$. Then $\{\Psi, \pi_\theta, \theta\}$ is an equilibrium with $\Psi = \{\psi^l, \psi^H\}$.

(ii) Let $\alpha \geq \dfrac{1}{2}$. Then $\max\{1, \frac{1}{4\alpha^2}\} = 1$ so $\beta > 1$ satisfies the condition on β in (b). In turn this ensures that $\frac{1-\alpha\sqrt{\beta}}{(1-\alpha)\sqrt{\beta}} < 1$. By (a) we have $\psi^l < \theta$. So pick $\theta \in (\frac{1-\alpha\sqrt{\beta}}{(1-\alpha)\sqrt{\beta}}, 1)$, and define π_θ as before, then $\{\Psi, \pi_\theta, \theta\}$ is an equilibrium with $\Psi = \{\psi^l, \psi^H\}$. ■

[9] An equivalent statement of (ii) is "There is an equilibrium such that $\theta = \psi^l = \frac{1}{\beta}$ and ψ^l is the only organizational focus if $\beta > 1$ and $\alpha < \frac{1}{1+\sqrt{\beta}}$"

Appendix B About perceptions

The model in the paper assumes that each agent in organization j perceives the social norm to be a weighted average of the "true" norm and the average action within his own organization. This gives rise to a uniform perception within each organization and hence a uniform action. The assumption that drives our results is that agents perceptions are biased; they assign greater weight to actions of agents within their own organization.

In actuality, each agent is likely to form his perception on a random sample of observations that arise from his individual interactions with other agents in the economy. The required bias arises from the fact that he interacts more often with agents within his own organization. In this appendix we argue that our conclusions go through with only small modifications as long as each agent's perception is based on a sufficiently large sample. The argument below is intuitive. To formalize it we would need to make specific assumptions about the distributions of x and y below, and the argument will only hold for specific classes of distributions.

Let actions across the economy be a random variable y with mean μ_y and variance σ_y^2, and actions within organization j be a random variable x with mean μ_x and variance σ_x^2. Let agent ij's perception be the average of n observations of actions, each drawn with probability α from organization j and probability $(1-\alpha)$ from the broader economy. For example, agent ij may make one observation each period (with the appropriate bias), with his current perception being the average of his most recent n observations. It follows immediately that the agent's perception is more stable if n is larger.

Then agent ij's sample consists of n draws of a random variable z with mean

$$\mu_z \;=\; \alpha\mu_x \;+\; (1-\alpha)\mu_y \tag{20}$$

and variance

$$\sigma_z^2 \;= \alpha^2\sigma_x^2 \;+\; (1-\alpha)^2\sigma_y^2 - \alpha(1-\alpha)\mathrm{Cov}(x,y)$$

Agent i's perception $\hat{\phi}_{ij}$ is the mean of his sample. Thus the perceptions of agents within organization j are draws of a random variable $\hat{\phi}_j$ distributed with mean μ_z and variance

$$\sigma_{\hat{\phi}_j}^2 = \frac{1}{n}\sigma_z^2.$$

It follows that if n is large, then there is little dispersion in perceptions within an organization.

Each agent takes an action $x_{ij} = \beta(\hat{\phi}_{ij})^2$. From the definition of variance it follows that the expectation of these actions is

$$\mu_x \;=\; \beta[(\mu_z)^2 + \sigma_{\hat{\phi}_j}^2] \;=\; \beta[(\mu_z)^2 + \frac{1}{n}\sigma_z^2]$$

Substitution in equation (20) we have

$$\mu_z \;=\; \alpha\beta(\mu_z)^2 + \alpha\beta\frac{1}{n}\sigma_z^2 \;+\; (1-\alpha)\mu_y. \tag{21}$$

Recall that $\mu_z = E(\hat{\phi}_j)$ is the mean of the organization's perceptions. The model in the paper assumes that all agents in j have the same perception ϕ_j. μ_y is the mean action in the economy, or θ in the notation of the model. If n is very large so that we can ignore the second term on the right-hand-side of (21), then equation (21) reduces to equation (14), and we recover the model in the paper. For finite n, equation (14) has an extra additive term $\alpha\beta\frac{1}{n}\sigma_z^2$, which yields solutions that are close to those for (14) if n is large.

In particular, consider a stable focus ϕ_j of the economy yielded by a solution for equation (14). Then there is a neighborhood $N(\phi_j)$ of ϕ_j such that a small displacement in perception will result in dynamics that correct this displacement (see Definition 1). If n is large enough, then the corresponding solution for μ_z in (21) is in the interior of $N(\phi_j)$. Further as $n \longrightarrow \infty$, $\mu_z \longrightarrow \phi_j$ and $\sigma^2_{\hat{\phi}_j} \longrightarrow 0$. Hence for any $p \in (0,1)$, there is n such that $\text{prob}\{\hat{\phi}_{ij} \notin N(\phi_j)\} < p$. Hence with probability $(1-p)^{N_j}$ the perceptions of all N_j agents in j remain close to ϕ_j. The arguments in the paper go through with this slightly modified conceptions of focus and stability.

Conversely, for any n and k we can calculate the probability that k or more members of j will develop perceptions outside $N(\phi_j)$. It seems a reasonable conjecture that the organization's "culture" will be disrupted if several of the agents have perceptions outside the stable neighborhood. This may have implications for anti-corruption policy. For example, a perception that "things are changing" as a result of anti-corruption initiatives may lead agents to disregard observations made in the past, and hence reduce the number of observations on which perceptions are based. This would make perceptions more volatile, and hence organizational cultures more prone to breaking down.