Corruption and the Hadleyburg effect⁰

Marco Celentani
Universidad Carlos III de Madrid

Juan-José Ganuza
Universitat Pompeu Fabra

April, 1999

⁰Address for correspondence: Marco Celentani, Department of Economics, Universidad Carlos III de Madrid, Getafe (Madrid) 28903, Spain; e-mail: celentan@eco.uc3m.es; fax +34-91-624 9875. Part of this research was done while Marco Celentani visited the Dipartimento di Teoria e Storia dell'Economia Pubblica of Università di Napoli Federico II under TMR fellowship ERBFMBICT960767 of the European Commission and Juan-José Ganuza visited the Institut d'Economie Industrielle of Université de Sciences Sociales de Toulouse and University of California, Los Angeles under Spanish Ministry of Education fellowship EX-50078539. We wish to gratefully acknowledge the hospitality of our former hosts as well as the financial support of our sponsors. Marco Celentani also acknowledges the financial support of grant ERBCHRXCT94-0458 of the European Commission and of grant UE95-0042 of the Spanish Ministry of Education.

We thank Alfredo Del Monte, Jean-Jacques Laffont, David K. Levine, Massimo Marrelli, Canice Prendergast, Patrick Rey, Manuel Santos, Jean Tirole and participants at the Learning and Social Institutions Conference in Alicante, Bellaterra Seminar at Universitat Autònoma de Barcelona, and Fourgeaud Seminar in Paris for useful discussions and suggestions. The usual disclaimer applies.

Abstract

We study the dynamics of corruption relying on two fundamental observations: (a) Given agents detected as corrupt are typically fired and replaced, the historical levels of corruption have an impact on current corruption through the distribution of corruption costs of old agents; (b) Institutions negatively affected by their agents' corrupt activities are likely to respond optimally to it thereby decreasing the payoff from being corrupt. We model this situation by considering an agent who is supposed to monitor a contractor's delivery of a product or service and can manipulate his report thus allowing the contractor to deliver lower quality in exchange for a bribe. Given the two generations of agents overlapping at any particular date, the administration sets an optimal level of quality in each period. We find that (i) A unique steady state level of corruption exists; (ii) Regardless of the initial distribution, apart from an initial period, equilibrium sequences are decreasing and converge to the steady state, a result we term the "Hadleyburg effect". We use these findings to study the dynamic response of corruption to both temporary and permanent shocks to the profitability of corruption and we find that the "Hadleyburg effect" has important positive and normative implications.

KEYWORDS: Corruption dynamics.

JEL classification numbers: C73, D73, K42.

“Why, you simple creatures, the weakest of all weak things is a virtue that has not been tested in the fire.” Mark Twain (1899) *The man that corrupted Hadleyburg*, reprint in Penguin Books, London, 1995, p. 46.

1 Introduction

Can the corruption opportunities created by privatization programs have a long lasting impact on the level of corruption? What are the long run consequences of a brief anti-corruption campaign? What is the dynamic response of corruption to permanent changes in the profitability of corruption? The objective of this paper is to contribute to the understanding of the dynamics of corruption and to ultimately be able to give answers to questions like the previous ones.

We consider a procurement problem¹ in which a principal plays against a continuum of agents who have a heterogenous component in their cost of becoming corrupt. Agents live for two periods and at each date the principal faces two different generations of agents, the young and the old. The model we develop is based upon three basic observations:

- A.1 Since agents who become corrupt have a relatively low corruption cost, the replacement of agents detected as corrupt with unknown agents, leads to the replacement of agents with a low expected cost of becoming corrupt with agents with a higher expected cost of becoming corrupt. This implies that the historical levels of corruption affect the distribution of the agents’ corruption costs.
- A.2 Institutions negatively affected by corruption respond optimally to it, reducing the payoff from being corrupt when corruption is expected to be widespread.
- A.3 People choose optimally the time at which they get corrupt trading off current payoff against the possible loss of future corruption opportunities resulting from discharge.

Our main result is that the historical levels of corruption may have a long lasting impact on corruption. Specifically, we show that a unique steady state exists and that the dynamics can be characterized as follows:

1. If the current level of corruption is higher than the steady state, corruption will monotonically converge down to it;

¹One of the first papers on corruption, Rose-Ackerman (1975), concentrated on public procurement and argued that this is a situation that lends itself to corrupt behavior as the fact that the government is often the only buyer (like in defense procurement contracts) and the fact that the product or service being bought are not homogeneous imply that little market information is available to monitor the decision making process.

2. Surprisingly, if the current level of corruption is lower than the steady state, corruption will first jump up and above the steady state and will then converge monotonically down to it.

We apply these results to study the dynamics following a number of different exogenous temporary or permanent shocks.

One of the implications we discuss is that a society that has not had valuable corruption opportunities in the past has failed to screen part of its corrupt members away and is therefore liable to experience sudden increases in the level of corruption if new corruption opportunities arise. This result is reminiscent of *The man that corrupted Hadleyburg* a tale in which Mark Twain (1899) relates how a man manages to induce an amazing surge in corruption among the citizens of Hadleyburg, a city known for his historical reputation for honesty. In the man's own words: "You were easy game. [...] As soon as I found out that you carefully and vigilantly kept yourself and your children *out of temptation*, I knew how to proceed. Why, you simple creatures, the weakest of all weak things is a virtue that has not been tested in the fire."² We see *The man that corrupted Hadleyburg* as a metaphor of the historical dependance of corruption and we have therefore termed our main result the "Hadleyburg effect".

Our paper is related to a number of other works in the area. In the model we propose the principal responds to the existence of corruption by distorting the procurement mechanism away from the first best without corruption. The idea that the possibility of corruption implies distortions in allocation mechanisms was first introduced by Banfield (1975). Laffont and Tirole (1991) formalized this intuition considering a simple procurement model in which the agent has to choose a supplier, and showing that the optimal mechanism may induce the principal to ignore quality altogether. In Celentani and Ganuza (1997) we generalized this procurement model along the lines of the literature on optimal procurement mechanisms and studied the equilibrium dynamics of corruption when agents who are detected as corrupt face a very high penalty for reiterated wrong doing.

The study of corruption in a dynamic setting was pioneered by Lui (1986) and Andvig and Moene (1990). Our initial setup is similar to the one introduced by Lui (1986) who analyzes two overlapping generations of public officials living for two periods. In Lui (1986) no principal plays an active role against corruption. He assumes that the amount of resources allocated to monitoring is given and that, because of agents' collusion, the detection probability is decreasing in the total number of corrupt agents. Given a fixed bribe, Lui (1986) analyzes the steady states of the system and how they respond to changes in the exogenous parameters, like the amount of resources allocated to monitoring, or corruption penalties. As will be clear from section 2, we assume that the detection probability is fixed, while the principal's procurement

²Twain (1899), p. 46. *Italic in the original.*

decisions are allowed to be dependent on total corruption, and we study equilibrium corruption allowing bribes to be endogenously determined.

Andvig and Moene (1990) consider a situation similar to the one described by Lui (1986) in that, while the monitoring probability is fixed, it is assumed that monitoring is performed by peers so that if a corrupt agent is overseen by a corrupt colleague, the former will bribe the latter into not giving him in, so that the expected penalty for being corrupt decreases with total corruption. Since the cost of finding a corrupt agent is lower the higher their number, private parties are also more willing to get involved in corrupt transaction the more agents they expect to be corrupt. As in Lui (1986), Andvig and Moene (1990) do not explicitly model a principal who plays any active role in the fight against corruption, apart from setting fixed monitoring probabilities and penalties, and their analysis is mainly focused on steady states.

More recently Tirole (1996) has shown how history can have an impact on equilibrium corruption through the interplay of individual and collective reputations of a set of agents. Tirole (1996) considers overlapping generations with a Poisson death and replacement process and studies a situation in which a principal has to decide whether to trust agents or not, depending on an imperfect record of their past behavior. If agents decide to be corrupt in one period, they get an extra payoff but face the possibility of losing the principal's future trust which implies lower future payoffs. As is clear, the threat of losing the principal's trust can have an impact only if the principal is expected to trust agents with spotless records in the future. If, on the other hand, the principal is expected to distrust all agents, regardless of their record, as is the case when the principal expects many of them to get corrupt, the threat of future losses as consequences of current corrupt behavior disappear, and the profitability of corruption increases. In this sense the profitability of corruption is independent of contemporaneous total corruption but is increasing in future total corruption.

Although the details of the models are very different, we think that our paper is most similar in spirit to Tirole (1996) as we explicitly model the role played by the principal and because we also focus on the way in which history affects equilibrium behavior.

A significant difference of this paper with respect to the papers mentioned above is that while they assume in a way or another that the profitability of corruption is increasing in (current or future) corruption we maintain that it is decreasing in contemporaneous corruption. As a consequence while all the papers mentioned above, exhibit in general the possibility of multiple steady states, in our model there will exist a unique steady state, so that we can account for long run differences in corruption only through different fundamentals and not through the selection of different equilibria or different initial conditions. Our main focus, on the other hand, will be on the dynamics of corruption away from the steady state.

The paper is structured in the following way. Section 2 presents the procurement model described above and studies the equilibrium in the static setting. To do this we first determine the optimal procurement mechanism chosen by a principal who believes that a given fraction of his agents are corrupt. We then study agents' corruption decisions given a procurement mechanism (set by the principal). Finally the equilibrium level of corruption and procurement mechanism is found for the static model. Section 3 introduces the dynamic setting which is the main focus of the paper and section 4 characterizes the equilibrium dynamics of corruption. Section 5 applies the result of the previous section and studies the dynamics of corruption following both temporary and permanent shocks. In section 6 we discuss our main findings. All proofs and an additional result are presented in the Appendices.

2 Corruption in a static setting

Consider a situation in which a principal (the administration) delegates an agent (public official) the purchase of a good or the realization of a project which can have different characteristics summarized in a quality level that enters the principal's utility function, $U^P(p, q) = V(q) - p$, where q is the quality of the project and p its price. The project is to be carried out by a firm whose cost of producing quality q is commonly known to be $C(q)$. The firm is assumed to be risk-neutral so that its payoff function is $U^F(p, q) = p - C(q)$ and its reservation level is normalized to zero. Since $C(q)$ is assumed to be common knowledge, the price which is paid to the firm for a given quality level q is the price that makes its profit equal to zero, $p = C(q)$.³

In the following we will consider payoff and cost functions that satisfy the following conditions: (a) $V(\cdot)$ is twice continuously differentiable, $V'(\cdot) > 0$, $V''(\cdot) < 0$, $\lim_{q \rightarrow 0^+} V'(q) = +\infty$, $\lim_{q \rightarrow +\infty} V'(q) = 0$; (b) $C(\cdot)$ is twice continuously differentiable; $C'(\cdot) > 0$, $C''(\cdot) \geq 0$. In order to give a more concrete flavor to some of the results, we will occasionally concentrate on the case in which $V(q) = \log q$ and $C(q) = \theta \cdot q$.

The agent is able to verify whether delivered quality coincides with the level specified in the procurement contract and has the option of misrepresenting his information and (illegally) allowing the firm to supply a lower level of quality, q_c , in exchange for a reward (bribe).

If an agent gets corrupt, he is detected with probability μ , and in this case he and the firm involved in the corrupt transaction are imposed penalties, B_A and B_F , respectively. The detection probability μ and the penalties B_A and B_F are thought of as features of the legal framework and in the following will therefore be assumed to be exogenously given. If an agent does not get corrupt, or if he does but he is not

³The purpose of looking at such a simplified setting is to avoid unnecessary complications, especially since the main focus of the paper is to look at the dynamics of corruption. The reader interested in a more realistic description of this kind of problem is referred to Celentani and Ganuza (1997) in which we assumed the existence of a finite number of potential contractors and asymmetric information on their cost functions. The model we present here retains all the relevant features of this more realistic setting while simplifying the derivation of the results.

detected, he is paid a fixed wage which is normalized to zero.

Agents are risk-neutral and are assumed to be heterogeneous in an additional component of the cost of being corrupt: an agent of type β has utility function

$$U^A(x, \beta) = \begin{cases} x & \text{if honest} \\ x - \beta & \text{if corrupt} \end{cases}$$

where x is expected income (including, possibly, any bribes paid by firms or expected penalties ensuing from corrupt activity) and β is the heterogenous cost component that can be given different interpretations, e.g., a moral cost (the disutility an agent derives from engaging in an illegal activity) or the cost of arranging a corrupt transaction. To simplify notation we will assume that β is uniformly distributed on $[0, 1]$. In the remainder of the paper we will always refer to β as the corruption cost, although it should be clear that it does not include the explicit cost of becoming corrupt associated with the possibility of being fined.

Summarizing, the time sequence of the model is the following:

1. Agents privately learn their type, β .
2. The principal announces the required quality level, q . Simultaneously agents decide whether to get corrupt or not. If an agent decides to be honest, he instructs the firm to produce the required quality level and reimburses it the cost of producing it, $C(q)$.⁴ If an agent decides to get corrupt, he bargains with the firm over a bribe b to be received in exchange for the agreement that the firm will be permitted to deliver quality level q_c rather than the one set by the principal while still receiving a cost reimbursement for the required quality level.
3. With probability μ the principal gets a perfect signal of delivered quality. If the delivered quality is less than the required one, the agent and the firm are imposed penalties B_A and B_F .

The next subsection will characterize the optimal quality level for a case in which a fixed fraction of the agents γ are known to be corrupt. Subsection 2.2 will then study how agents take a rational decision to get corrupt in the face of a given quality level required by the principal. Subsection 2.3 provides the equilibrium levels of corruption and quality in a static setting.

2.1 The optimal quality requirement with corrupt agents

Suppose the principal sets a quality standard q . In this case an honest agent guarantees that the delivered quality is the required one and pays the firm a price equal to the cost of producing it, $C(q)$. A corrupt agent requires only quality q_c of the firm but announces that the delivered quality is q and accordingly

⁴Recall that the firm's reservation level is assumed to be 0.

pays $C(q)$ to the firm. Given this, when the principal believes that with probability γ the agent is corrupt, his maximization problem will be

$$\max_q (1 - \gamma) (V(q) - C(q)) + \gamma (V(q_c) - C(q)) = (1 - \gamma) V(q) + \gamma V(q_c) - C(q)$$

and the optimal quality level given a probability γ for the agent to be corrupt, $q(\gamma)$, is given by the following first-order condition:

$$(1 - \gamma) V'(q(\gamma)) = C'(q(\gamma)) \quad (1)$$

equating the expected marginal value of quality to its marginal cost. It is easy to see that, under the assumptions specified above, $q(\gamma)$ exists and is unique and that

$$q'(\gamma) = \frac{V'(q^*)}{(1 - \gamma) V''(q^*) - C''(q^*)} < 0$$

so that the optimal quality level is decreasing in the probability of the agent being corrupt. In this simplified setting this result recalls Banfield's (1975) view that "[N]arrowing discretion [...] while preventing the agent from doing (corrupt) things that are slightly injurious to the principal it may at the same time prevent him from doing (non-corrupt) ones that would be very beneficial to him. If simply to prevent corruption an agent is given a narrower discretion than would be optimal if there were no corruption, whatever losses are occasioned by his having a sub-optimal breadth of discretion must be counted as costs of preventing corruption."⁵

It is probably important to clarify at this point that, as was spelled out in the beginning of this section, we are describing a situation in which the principal and the agents move simultaneously so that in equilibrium the principal makes a quality decision while taking the agents' decision to be corrupt as given. This is to say that we are not allowing the principal to set his quality standard in advance, thus endowing him with the ability to commit to a quality level which is observed by the agents before taking their decision to be corrupt or not. With this set-up the reason why the principal sets a lower quality standard with corruption is, as shown by equation (1), that the expected marginal value of quality q is decreasing in the probability of the agent being corrupt, γ , as quality q is only delivered with probability $1 - \gamma$.

It is easy to show that if the principal could observably commit to a quality level before the agents decided whether to be corrupt or not, the equilibrium levels of corruption and quality would both be lower but that all the dynamics results to be discussed below would be confirmed.

⁵Banfield (1975), p. 590.

2.2 The decision to be corrupt

The goal of this subsection is to study how the payoff from being corrupt for an agent depends on the quality standard set by the principal and how an agent decides whether to be corrupt or not given such quality standard.

If an agent decides to get corrupt, he bargains with the firm over the bribe to be received in exchange for the fact that the firm will be paid $C(q)$ but will be asked to deliver only quality $q_c < q$. We assume that the outcome of the bargaining process between the corrupt agent and the firm is the solution of a generalized Nash bargaining problem in which the firm knows the agent's idiosyncratic corruption cost β .⁶ This assumption is only meant to simplify the calculation of the equilibrium in the intertemporal setting to be introduced and is by no means essential for our results.

Let $\pi_A(\beta)$ denote the agreement payoff to a (corrupt) agent with corruption cost β , and d_A his disagreement point. Let $\pi_F(q)$ denote the agreement payoff from a corrupt transaction to the firm when the quality standard set by the principal is q and let d_F denote the firm's disagreement point. The bribe b will then be given by the solution to the following problem, where its dependence on β and q is left implicit:

$$\max_b (\pi_A - d_A)^\alpha (\pi_F - d_F)^{1-\alpha}$$

A type β corrupt agent's agreement payoff is

$$\pi_A = b - \mu B_A - \beta, \tag{2}$$

i.e., bribe minus expected penalty, minus corruption cost; his disagreement level will simply be his constant reservation wage which has been normalized to zero, $d_A = 0$. Given quality standard q , the firm's agreement payoff will be

$$\pi_F(q) = C(q) - C(q_c) - b - \mu B_F$$

i.e., the payment it receives, minus the cost of producing quality q_c , minus the bribe, minus the expected penalty; its disagreement point will be its payoff without corruption, $d_F = 0$. Given this, it is easy to show that the equilibrium bribe paid by the firm to the agent is

$$b = \alpha [C(q) - C(q_c) - \mu B_F] + (1 - \alpha) [\beta + \mu B_A] \tag{3}$$

and is therefore increasing in the agent's bargaining power α .⁷

We now want to compute the probability that an agent decides to get corrupt when the principal sets quality standard q . To do this we need to compute the probability that, for a given q , the corruption cost

⁶ An alternative equivalent formulation is the one pursued in Celentani and Ganuza (1997) in which the matching cost β is a sunk cost that does not enter the bargaining problem.

⁷ Notice that $C(q) - C(q_c) - \mu B_F - \beta - \mu B_A > 0$ is a necessary condition for an agreement to take place and this implies that $C(q) - C(q_c) - \mu B_F > \beta + \mu B_A$, which in turn implies that b is increasing in α .

of an agent is sufficiently low for him to get corrupt, or equivalently the type $\beta(q)$ who is exactly indifferent between being honest and corrupt. Straightforward calculations show that

$$\beta(q) = C(q) - C(q_c) - \mu(B_A + B_F)$$

or, letting $k = C(q_c) + \mu(B_A + B_F)$ denote the non-idiosyncratic cost of agreeing to (illegally) deliver quality q_c ,

$$\beta(q) = C(q) - k \tag{4}$$

Given β has been assumed to be uniformly distributed on $[0, 1]$, without loss of generality, in the following we will always assume that the type of agent who is exactly indifferent between being corrupt or not is less than or equal to 1. Under this assumption the mass of corrupt agents will be exactly the type of agent who is indifferent between being honest and corrupt

$$\gamma(q) = \beta(q). \tag{5}$$

Since $\beta'(\cdot) > 0$, it is easy to see that the mass of corrupt agents is increasing in the desired quality level set by the principal.

2.3 Equilibrium in a static setting

The equilibrium of the static setting is given by equations (1) and (5) representing the best response of the principal to the aggregate strategy of the agents and the aggregate best response of the agents to the strategy of the principal, respectively. Substituting $q(\gamma)$ into (4) and (4) into (5) we get

$$\gamma = C(q(\gamma)) - k. \tag{6}$$

Let

$$\Gamma(\gamma) = C(q(\gamma)) - k. \tag{7}$$

Since $\Gamma(\gamma)$ denotes the expected payoff from being corrupt (gross of the agents' heterogeneous corruption costs) when corruption is equal to γ and, therefore, when the principal requires quality $q(\gamma)$, in the following we will refer to it as the profitability of corruption. Since all agent with a corruption cost less than or equal to the profitability of corruption decide to be corrupt, from (6) and (7) the equilibrium level of corruption, $\hat{\gamma}$, is defined by the following equation

$$\hat{\gamma} = \Gamma(\hat{\gamma}). \tag{8}$$

Since $C'(q(\gamma)) > 0$ and $q'(\gamma) < 0$, $\Gamma'(\gamma) = C'(q(\gamma))q'(\gamma) < 0$. Since $q(1) = 0$, $\Gamma(1) = -k < 0$ and a sufficient condition for an interior equilibrium to exist is $\Gamma(0) = C(q(0)) - k > 0$.

In the case in which $V(q) = \log q$ and $C(q) = \theta q$, from equations (1) and (5)

$$\begin{aligned} q(\gamma) &= \frac{1-\gamma}{\theta} \\ \gamma(q) &= \theta q - k. \end{aligned}$$

and the equilibrium of the static setting, provided it is interior, will be (see figure 1):

$$\begin{aligned} \hat{q} &= \frac{1+k}{2\theta} \\ \hat{\gamma} &= \frac{1-k}{2}. \end{aligned}$$

The static model presented in this section could be used for comparative statics exercises to study the consequences of both temporary or permanent shocks to the profitability of corruption. As one could expect, all changes implying a higher (lower) profitability of corruption, such as changes in the penalties, B^A and B^F , or in the probability of detection, μ , would lead to higher (lower) levels of corruption.

The purpose of the rest of the paper is to show that the dynamics of corruption can be very different from the intuition gathered by such exercises if different generations of agents overlap.

3 Intertemporal optimization and the dynamics of corruption

Suppose that at each date $t = 0, 1, 2, \dots$ two generations of agents exist, the young and the old, each living for two periods and each having unit mass. While all generations born from date $t = 0$ on will be assumed to be identical at birth, in that they have the same preferences and corruption costs drawn from a uniform distribution over the interval $[0, 1]$, the two generations that overlap at any date $t = 1, 2, \dots$, differ in the following respects:

1. Since some of the old were corrupt when young and with probability μ were detected and replaced,⁸ the probability distribution of the old agents' corruption costs, β , will be different from the young agents' corresponding distribution.
2. Since agents live for two periods, old agents do not care about the future whereas young agents do so long as their discount factor is strictly positive. Even though all the results of the paper hold for all discount factors (even 0),⁹ to simplify notation, we will assume that the young agents' discount factor is equal to 1.

⁸The replacement of an agent detected as corrupt by another agent is discussed below.

⁹Since the qualitative results of the paper hold also in the case in which the discount factor is equal to 0, they in fact do not depend on young agents being concerned about the future. We thank Canice Prendergast for making us recognize that. The relationship between the case in which young agents have a 0 discount factor and the case in which they care about future payoffs is discussed in subsection 4.1.

Since we will be interested in equilibrium sequences following exogenous shocks, in the following we will consider the possibility that the generation of old agents alive at $t = 0$ has corruption costs drawn from an arbitrary distribution over the interval $[0, 1]$.

We assume that the principal lives one period only, or equivalently that his discount factor is equal to 0. While simplifying the analysis, this assumption seems reasonable as the tenure of a principal (a top manager in a firm, a politically appointed top official in a bureaucracy, or an elected government in a democracy) is usually much shorter than lower ranked officials' tenures.

Since young agents will be alive tomorrow while old will not, their decisions will be different. Since old agents have no future, we can use equation (4) to determine the type of old agent in t who will be indifferent between being corrupt or not, β_t^O , given the quality standard at t , q_t (the dependence of β_t^O on q_t will be left implicit in the following):

$$\beta_t^O = C(q_t) - k. \quad (9)$$

As in subsection 2.2 we will assume without loss of generality that $\beta_t^O \leq 1$.

Young agents may decide to get corrupt in one, two periods, or none at all. Since, however, an agent who is detected as corrupt when young is known to have a low corruption cost, it is reasonable to assume that he is fired and replaced by another agent who is drawn from the prior distribution on $[0, 1]$ and has therefore a higher expected corruption cost.¹⁰ As was remarked above, the fact that an agent caught as corrupt is replaced by another agent has an impact on the distribution of future corruption costs.

The dependence of the distribution of corruption costs on the past can be expected to have important implications on equilibrium dynamics of corruption and the rest of the paper will pursue this argument.

The decision of a young agent in t is determined by the expected profitability of corruption in t and in $t + 1$ which are ultimately determined by the (equilibrium) quality standards in these periods, q_t and q_{t+1} . Since we want to incorporate the agents' consistent expectations into the decision making, we will consider sequential equilibria.

Let $\beta_t^Y(q_t, q_{t+1})$ denote the type of young agent in t who is exactly indifferent between being corrupt or not in t given quality levels q_t and q_{t+1} . Straightforward although tedious calculations in Appendix A.1 show that for $t \geq 0$

$$\beta_t^Y(q_t, q_{t+1}) = \begin{cases} \frac{C(q_t) - \mu\alpha C(q_{t+1})}{1 - \mu\alpha} - k & \text{if } q_t < q_{t+1} \\ C(q_t) - k & \text{if } q_t \geq q_{t+1}. \end{cases} \quad (10)$$

Since the profitability of corruption is increasing in q , equation (10) makes it clear that if future quality is less than the current one, $q_t \geq q_{t+1}$, the young agent who is indifferent between being corrupt or not in t

¹⁰This replacement mechanism is much more compelling once we recognize that, as it will be shown below, in equilibrium an agent who got corrupt when young will become corrupt with probability 1 when old.

will not be corrupt in $t + 1$. This means that he does not care about future corruption opportunities and therefore has the same corruption cost as the indifferent old, $\beta_t^Y(q_t, q_{t+1}) = C(q_t) - k$. By the same token, however, if $q_t < q_{t+1}$, a young agent who is indifferent between being corrupt or not in t does care about future corruption opportunities and may therefore be indifferent between being corrupt or not in t only if his corruption cost is sufficiently lower than the one of the indifferent old, $\beta_t^Y(q_t, q_{t+1}) = \frac{C(q_t) - \mu\alpha C(q_{t+1})}{1 - \mu\alpha} - k < C(q_t) - k$.

Under the assumption that $\beta_t^Y(q_t, q_{t+1}) \leq 1$ and recalling that the distribution of β has been assumed to be uniform over the interval $[0, 1]$, the fraction of young agents who get corrupt at date t is

$$\gamma_t^Y = \beta_t^Y(q_t, q_{t+1}). \quad (11)$$

To compute the mass of old agents who get corrupt in $t \geq 1$, it is necessary to keep into account that some of them already got corrupt in $t - 1$, and a fraction μ of those were caught and, having been fired, will not be able to get corrupt in t . On the other hand since the agents who were detected as corrupt are replaced by “new” old agents drawn from the prior distribution, we need to keep into account that also some of the newly hired agents will find profitable to be corrupt.

Consider the distribution of corruption costs of a generation of old agents such that all types below β' were corrupt when young and a fraction μ of them was caught and fired and has been replaced by “new” old agents (who will live for only one period¹¹) drawn from the prior probability distribution, which is the uniform on $[0, 1]$. On one side this implies that, because of the firings, the density of agents with corruption cost below β' is reduced by μ . On the other hand since $\mu\beta'$ agents were fired and the same mass of agents is hired in replacement, and since the substitutes are drawn from uniform distribution on $[0, 1]$, this implies that density $\mu\beta'$ has to be added for all $\beta \in [0, 1]$. The new distribution can then be characterized by the following density function, where $\beta' \in [0, 1]$ and $\mu \in (0, 1)$

$$f(\beta; \beta', \mu) = \begin{cases} 1 - \mu + \mu\beta' & \beta \in [0, \beta') \\ 1 + \mu\beta' & \beta \in [\beta', 1] \end{cases} \quad (12)$$

represented in Figure 2, or, equivalently, by the associated distribution function

$$F(\beta; \beta', \mu) = \int_0^\beta f(s; \beta', \mu) ds = \begin{cases} (1 - \mu + \mu\beta')\beta & \beta \in [0, \beta') \\ (1 - \mu + \mu\beta')\beta' + (1 + \mu\beta')(\beta - \beta') & \beta \in [\beta', 1] \end{cases} \quad (13)$$

represented in Figure 3.

Given this, it is easy to see that, given β_{t-1}^Y and β_t^O , the mass of old agents who get corrupt in $t \geq 1$ is

$$\gamma_t^O = F(\beta_t^O; \beta_{t-1}^Y, \mu). \quad (14)$$

¹¹It is interesting to know that, according to a study of The Economist with the Enrico Mattei Foundation on the 100 most powerful posts in Italy, despite of the fact that in the 5 years following the 1992 Clean Hands investigation campaign the turnover rate has been of 71%, up from 54% in each of the two five-year periods preceding 1992, the average age of people holding those posts went down to just above 59 years, only six months less than 5 years before; *The Economist* (1997).

The reason why t is constrained to be greater than or equal to 1 is that the equation (14) provides the mass of old agents who get corrupt at t using the distribution of old agents' corruption costs derived from the following assumptions: 1) That the distribution of their corruption costs at birth was the uniform over $[0, 1]$; 2) That all types below β_{t-1}^Y got corrupt at the previous date. At $t = 0$ on the other hand we want to consider the case of an arbitrary distribution of old agents' corruption costs and this implies that equation (14) is only applicable for $t \geq 1$. Since it is easy to show that $\beta_{t-1}^Y < \beta_t^O$ if and only if $q_{t-1} < q_t$, from (14) we have

$$\gamma_t^O = \begin{cases} \beta_t^O - \mu\beta_{t-1}^Y + \mu\beta_{t-1}^Y\beta_t^O & \text{if } q_{t-1} < q_t \\ (1 - \mu)\beta_t^O + \mu\beta_{t-1}^Y\beta_t^O & \text{if } q_{t-1} \geq q_t. \end{cases} \quad (15)$$

The following section turns to the characterization of equilibrium sequences.

4 Equilibrium sequences of corruption

The previous results have shown how the fraction of old and young agents who decide to get corrupt depends on the dynamics of quality indices. Appendix A.2 shows that given q_{t-1} , q_t , and q_{t+1} , corruption at date $t \geq 1$ ¹² will be

$$\begin{aligned} \gamma_t = & \frac{1}{2} [(2 - \mu)(C(q_t) - k) + \mu(C(q_{t-1}) - k)(C(q_t) - k) \\ & - \delta_F \frac{\mu\alpha}{1 - \mu\alpha} (C(q_{t+1}) - C(q_t)) \\ & - \delta_P \frac{\mu}{1 - \mu\alpha} (C(q_{t-1}) - C(q_t))(1 - \mu\alpha(C(q_t) - k))] \end{aligned} \quad (16)$$

where δ_F and δ_P are indicator functions depending respectively on the future and past dynamics of corruption: $\delta_F = 1$ if $q_t < q_{t+1}$, and $\delta_F = 0$ otherwise; $\delta_P = 1$ if $q_{t-1} < q_t$, and $\delta_P = 0$ otherwise.

Equation (16) can be interpreted as follows: The first term gives a level of corruption that depends on current quality standard and while it also depends on the previous period (absolute) quality standard, it is independent of whether quality standards have *increased* or *decreased*. Realized corruption, however, does depend on whether quality standards increase or decrease over time, and this can add or subtract mass of corrupt agents. When $q_t < q_{t+1}$, young agents foresee that their opportunities to get corrupt are more valuable in the future than they are now, since higher future quality standards imply that the equilibrium bribe will be higher; For this reason some young agents will refrain from getting corrupt in anticipation of a higher expected payoff tomorrow: In this case $\delta_F = 1$ and since $\frac{\mu\alpha}{1 - \mu\alpha} (C(q_{t+1}) - C(q_t)) > 0$ a fraction of corrupt agents is to be subtracted. When $q_{t-1} < q_t$, the indifferent young agent in $t - 1$ had a lower corruption cost than the indifferent old in t both because of the direct effect of a lower quality q_{t-1} (that implies a lower profitability of corruption) and because young agents anticipated that the opportunity to

¹²To see why $t \geq 1$, refer to the discussion of equation (14).

be corrupt would be more valuable when old in t . This means that a lower fraction of old agents was caught when young which in turn implies that the average corruption cost of old agents is lower: In this case $\delta_P = 1$ and since $\frac{\mu}{1-\mu\alpha} (C(q_{t-1}) - C(q_t)) (1 - \mu\alpha(C(q_t) - k)) < 0$ a fraction of agents is added to the mass of corruption determined by the current level of discretion.

Now recall that we made the assumption that the principal lived for one period only, or, equivalently, that he was completely myopic. Substituting the optimal quality standard for the static setting $q_t = q(\gamma_t)$ into (16) and using (7) we get the following second order difference equation in γ , for $t \geq 1$ ¹³:

$$\begin{aligned} \gamma_t = & \frac{1}{2} [(2 - \mu)\Gamma(\gamma_t) + \mu\Gamma(\gamma_{t-1})\Gamma(\gamma_t) \\ & - \delta_F \frac{\mu\alpha}{1 - \mu\alpha} (\Gamma(\gamma_{t+1}) - \Gamma(\gamma_t)) \\ & - \delta_P \frac{\mu}{1 - \mu\alpha} (\Gamma(\gamma_{t-1}) - \Gamma(\gamma_t)) (1 - \mu\alpha\Gamma(\gamma_t))] \end{aligned} \quad (17)$$

where $\delta_F = 1$ if $\gamma_t > \gamma_{t+1}$, and $\delta_F = 0$ otherwise, $\delta_P = 1$ if $\gamma_{t-1} > \gamma_t$, and $\delta_P = 0$ otherwise. Using the definitions of equations (32), (33), and (34) in Appendix A.4, we can rewrite equation (17) as

$$E(\gamma_t, \gamma_{t-1}) + \delta_F E_F(\gamma_t, \gamma_{t+1}) + \delta_P E_P(\gamma_t, \gamma_{t-1}) = 0 \quad (18)$$

where $\delta_F = 1$ if $\gamma_t > \gamma_{t+1}$, and $\delta_F = 0$ otherwise, $\delta_P = 1$ if $\gamma_{t-1} > \gamma_t$, and $\delta_P = 0$ otherwise.

The following proposition states that there is a unique steady state of the equilibrium system, characterizes it, and provides its comparative statics.

Proposition 1 (1) *There exists only one steady state of the equilibrium system, γ^* . (2) Let β^{Y*} and β^{O*} be the corruption cost that makes, respectively, a young and an old agent indifferent between being corrupt or not in the steady state. Then, $\beta^{Y*} = \beta^{O*} = \beta^* = \Gamma(\gamma^*)$. (3) The steady state is independent of the agents' bargaining power, α , and decreasing in the detection probability, μ , penalties B_A and B_F , and quality delivered under corruption, q_c .*

Proof. Appendix A.3.

While parts 1 and 3 of Proposition 1 are fairly intuitive, it is probably worth commenting part 2. Part 2 shows that in the steady state the corruption cost that makes agents indifferent between being corrupt or not is the same for old and young agents. The reason why this is true is that in the steady state the expected profitability of corruption opportunities is constant over time and an agent whose corruption cost is such that he is indifferent when old between being corrupt or not, when young attaches zero value to future corruption opportunities and therefore only cares about current payoffs—just like when he himself will be old.

¹³To see why $t \geq 1$, again, refer to the discussion of equation (14).

Next proposition provides a complete characterization of the dynamics of equilibrium sequences of corruption starting from an initial distribution of old agents' corruption costs.

Proposition 2 *Let the initial distribution of old agents at $t = 0$ be $G(\beta)$, a continuous function, and let $\{\gamma'_t\}_{t=0}^\infty$ denote the equilibrium sequence starting from this distribution.*

1. *If*

$$G(\beta^*) = F(\beta^*; \beta^*, \mu) \quad (19)$$

then there exists a unique equilibrium sequence $\{\gamma'_t\}_{t=0}^\infty$, with $\gamma'_t = \gamma^$, $t = 0, 1, \dots$*

2. *If*

$$G(\beta^*) > F(\beta^*; \beta^*, \mu) \quad (20)$$

then there exists a unique equilibrium sequence $\{\gamma'_t\}_{t=0}^\infty$, with $\gamma'_0 > \gamma'_1 > \gamma'_2 > \dots > \gamma^$.*

3. *If*

$$G(\beta^*) < F(\beta^*; \beta^*, \mu) \quad (21)$$

then there exists a unique equilibrium sequence $\{\gamma'_t\}_{t=0}^\infty$, with $\gamma'_0 < \gamma^$, $\gamma'_1 > \gamma'_2 > \dots > \gamma^*$.*

Proof. Appendix A.4.

Rather than presenting the formal proof of Proposition 2, which is provided in Appendix A.4, the following subsection will discuss the intuition behind it.

4.1 Discussion

Proposition 2 characterizes equilibrium sequences starting from an arbitrary distribution of old agents' corruption costs described by its distribution function. Before discussing the intuition of Proposition 2 it is important to notice that, given agents are heterogeneous with respect to their corruption costs, the appropriate initial condition is not a level of corruption, but the distribution of old agents' corruption costs, as the former is a sufficient statistics of the latter only in equilibrium sequences.

Both to simplify the arguments to be presented and to show that they do not depend on the fact that young agents care about the future, in what follows we will refer to the case in which agents have a 0 discount factor. In this case young agents act as old agents and the type of young agent who will be indifferent between being corrupt or not is simply $\beta_t^Y = \Gamma(\gamma_t)$. After completing the argument for this case we will argue that the same results have to hold when young agents care about the future and that there is a sense in which this reinforces the argument to be discussed next.

Part 1 of Proposition 2 is fairly intuitive. If the initial distribution $G(\cdot)$ is such that old agents whose corruption costs are less than the profitability of corruption in the steady state, $\beta^* = \Gamma(\gamma^*)$, have the same mass as in the steady state ($G(\beta^*) = F(\beta^*; \beta^*, \mu)$), then in the unique equilibrium sequence corruption will always be equal to the steady state.

Consider now the level of corruption at $t = 0$ when the initial distribution, $G(\beta)$, is such that $G(\beta^*) \neq F(\beta^*; \beta^*, \mu)$. Parts 2 and 3 of Proposition 2 state the following:

- (C1) If the mass of agents with a corruption cost lower than that of the indifferent agent in the steady state is *higher* than the corresponding mass in the steady state distribution, $G(\beta^*) > F(\beta^*; \beta^*, \mu)$, the level of corruption in $t = 0$ will be *higher* than in the steady state, $\gamma'_0 > \gamma^*$.
- (C2) If the mass of agents with a corruption cost lower than that of the indifferent agent in the steady state is *lower* than the corresponding mass in the steady state distribution, $G(\beta^*) < F(\beta^*; \beta^*, \mu)$, the level of corruption in $t = 0$ will be *lower* than in the steady state, $\gamma'_0 < \gamma^*$.

To see the intuition behind the previous statements notice that when agents have a 0 discount factor, their decisions at any date are independent of the future. In the steady state, as was stated in Proposition 1 we have that $\beta^{Y*} = \beta^{O*} = \beta^* = \Gamma(\gamma^*)$. This means that in every period a mass $\mu\beta^*$ of young agents are detected as corrupt and fired, and are replaced in the following period by an equal mass of old agents who are drawn from the prior probability distribution. The combination of these two effects determines the steady state distribution of old agents' corruption costs, $F(\beta; \beta^*, \mu)$.

Notice also that at any date the level of corruption is determined by two elements: (1) The corruption cost of the agent who is indifferent between being corrupt or not (when agents are myopic, this indifference corruption cost is the same for both the young and the old agents and is equal to the profitability of corruption, $\Gamma(\gamma_t)$); (2) The distribution of old agents' corruption costs.

Consider now case (C1) and suppose that, counter to the claim, the level of corruption is below the steady state. If corruption is below the steady state, the profitability of corruption is higher and so is the indifference corruption costs for both the young and the old. This implies that for corruption to be below the steady state, a necessary condition is that the mass of old agents with corruption costs below the steady state indifferent agent corruption cost is lower than the corresponding mass for the steady state distribution. Since this is just the opposite of the condition in case (C1) a contradiction is obtained and the level of corruption is seen to be higher than in the steady state. A parallel argument explains why under the condition of case (C2) the initial level of corruption has to be lower than in the steady state.

Having characterized corruption in $t = 0$, we now move on to the characterization of equilibrium in $t = 1, 2, \dots$. Consider the following questions:

- Will an initial level of corruption *above* the steady state ($\gamma'_0 > \gamma^*$) generate a distribution of old agents' corruption costs satisfying condition (C1) or (C2) in $t = 1$?
- Will an initial level of corruption *below* the steady state ($\gamma'_0 < \gamma^*$) generate a distribution of old agents' corruption costs satisfying condition (C1) or (C2) in $t = 1$?

The replacement of corrupt agents, surprisingly, implies that the answer to both questions is (C1), so that the level of corruption from the second period on will be higher than in the steady state. To see the intuition behind this result consider the following arguments.

Suppose corruption is above the steady state. This means that, given the profitability of corruption is below its steady state level, some agents who would be corrupt in the steady state in fact choose not to be corrupt. This means that, as compared with the steady state, when corruption is above the steady state, the replacement of some agents with corruption costs below the steady state indifference level with agents whose corruption costs can be both below or above the steady state indifference level fails to occur. This in turn implies that the mass of old agents with corruption costs less than or equal to the steady state indifference level is higher than the corresponding mass in the steady state distribution, and a level of corruption above the steady state generates a distribution of old agents' corruption costs that satisfies condition (C1).

Suppose now corruption is below the steady state. This means that, given the profitability of corruption is above its steady state level, some agents who would not be corrupt in the steady state in fact choose to be corrupt. This means that, as compared with the steady state, when corruption is below the steady state, some agents with corruption costs above the steady state indifference level are replaced with agents whose corruption costs can be both below or above the steady state indifference. This in turn implies that the mass of old agents with corruption costs less than or equal to the steady state indifference level is higher than the corresponding mass in the steady state distribution, so that a level of corruption below the steady state generates a distribution of old agents' corruption costs that, again, satisfies condition (C1).

Given the above arguments can be reiterated we now know that apart from the initial period, $t = 0$, corruption in all subsequent periods will be above the steady state.

Parts 2 and 3 of Proposition 2 also claim that corruption in periods $t = 1, 2, \dots$ is strictly decreasing. To see why suppose, contrary to the claim, that there is a period $t = 1, 2, \dots$ such that $\gamma_t < \gamma_{t+1}$. Recall that by definition

$$\gamma_{t+1} = \frac{1}{2} (\gamma_{t+1}^O + \gamma_{t+1}^Y) \quad (22)$$

and by (11)

$$\gamma_{t+1}^Y = \beta_{t+1}^Y; \quad (23)$$

Finally, recall from (15) that when corruption is increasing, $\gamma_t < \gamma_{t+1}$,

$$\gamma_{t+1}^O = (1 - \mu) \beta_{t+1}^O + \mu \beta_t^Y \beta_{t+1}^O. \quad (24)$$

From (22), (23) and (24) we then have that, if $\gamma_t < \gamma_{t+1}$:

$$\gamma_{t+1} = \frac{1}{2} \left((1 - \mu) \beta_{t+1}^O + \mu \beta_t^Y \beta_{t+1}^O + \beta_{t+1}^Y \right). \quad (25)$$

In a similar way we also get

$$\gamma^* = \frac{1}{2} \left((1 - \mu) \beta^* + \mu \beta^* \beta^* + \beta^* \right). \quad (26)$$

Since we know from above that $\gamma_t > \gamma^*$, for $t = 1, 2, \dots$, we have $\beta_{t+1}^O < \beta^*$, $\beta_t^Y < \beta^*$, and $\beta_{t+1}^Y < \beta^*$ and it is easy to see that equations (25) and (26) are inconsistent with $\gamma_{t+1} > \gamma^*$. This shows that no $t = 1, 2, \dots$ can exist such that $\gamma_t < \gamma_{t+1}$ for $t = 1, 2, \dots$, so that equilibrium sequences have to be decreasing and therefore converging to the steady state.

We call the result of Proposition 2 the “Hadleyburg effect” since it stresses the dynamic implications of the historical levels of corruption exemplified in *The Man that corrupted Hadleyburg* that relates how the past absence of corruption opportunities may make a society vulnerable to current corruption opportunities as stated in condition (C2). The closing of the tale summarizes how the citizens of Hadleyburg became aware of the dynamic implications of the lack of corruption opportunities: “By act of the Legislature—upon prayer and petition—Hadleyburg was allowed to change its name to (never mind what—I will not give it away), and leave one word out of the motto that for many generations had graced the town’s official seal.” [See figure 5]¹⁴

As compared with the case in which their discount factor is equal to 0, when young agents care about the future, they adjust their choices only if future corruption opportunities are more profitable than current ones, i.e., only if future corruption is anticipated to be lower than current. This is taken into account in equation (18) by the function $\delta_F E_F(\gamma_t, \gamma_{t+1})$ which is nonzero only if $\gamma_t > \gamma_{t+1}$.

Given the results discussed above, we can now ask the following question: Do the qualitative features of equilibrium sequences for zero discount factor case change when young agents care about the future? In particular, can equilibrium sequences in the latter case be increasing?

The answer to this question is no, as, if they were increasing, although young agents care about the future, their anticipation of higher corruption and therefore of less profitable corruption opportunities would make them disregard the future and act as if they were myopic. In formal terms, this would imply $\delta_F = 0$ and equation (18) would be the same as in the case in which agents are not forward looking for

¹⁴Twain (1899), page 60.

which equilibrium sequences have been shown to be decreasing. Similar arguments can be used to rule out nonmonotonic sequences¹⁵ and the result is that, away from the steady state, equilibrium sequences have to be strictly decreasing which implies that they have to converge to the steady state from above.

Since corruption is decreasing, the dynamics of the system is described by equation (18) with $\delta_F = \delta_P = 1$. Given in this case corruption at any date t depends both on the historical level of corruption and on the expectation on its future levels, it seems in principle possible that equilibrium sequences may be not unique as a lower level of corruption at date t may result if corruption at date $t + 1$ is also expected to be lower. The proof of Proposition 2 shows that this is not possible as for two different levels of corruption at date t to belong to different equilibrium sequences, it is seen that the differences of the elements of the continuation sequences have to be increasing. Since this is inconsistent with the fact that both continuation sequences tend to the steady state, equilibrium sequences are shown to be globally unique.

To compare the case in which young agents do not care about future payoffs with the case in which they do, notice that in the former case they myopically ignore that corruption opportunities are more valuable in the future as corruption decreases over time. For this reason, young agents are more prone to corruption when they do not care about the future and, given an initial distribution $G(\beta)$ with $G(\beta^*) > F(\beta^*; \beta^*, \mu)$, corruption will unambiguously be higher in the first period.

By the same token, when agents do care about future payoffs, the previous argument implies that the distribution of old agents in the following period will have a higher mass of agents with corruption costs below the steady state level, since more agents with corruption costs below the steady state indifference level have refrained from being corrupt in anticipation of future more valuable corruption opportunities. This effect may in fact more than compensate the lower corruption propensity of young agents in the following period, so that corruption in the following period may be higher than in the case in which agents are myopic. For this reason it is possible that, while the starting point of the equilibrium sequence is higher when agents are myopic than when they care about future payoffs, convergence in the latter case may be slower so that there exists a τ such that corruption is higher for all $t > \tau$ in the latter case.

When $G(\beta^*) < F(\beta^*; \beta^*, \mu)$, on the other hand the level of corruption in $t = 0$ will be the same in both cases and will then be higher in $t = 1$ in the case of myopic agents. The same arguments discussed above may be applied for the continuation sequences starting from $t = 2$.

4.2 Additional characterization of equilibrium sequences

The following Proposition provides an additional characterization of equilibrium sequences.

¹⁵When young agents care about the future, equilibrium sequences are characterized by second order difference equations. For this reason, to conclude that equilibrium sequences are decreasing, it is necessary to also rule out nonmonotonic subsequences. The Appendix provides a formal proof of this result.

Proposition 3 (1) Consider two different equilibrium sequences $\{\gamma_t\}$ and $\{\gamma'_t\}$, each associated to a given initial distribution. Suppose that $\gamma_\tau \geq \gamma^*$, $\gamma'_\tau > \gamma^*$, and $\gamma_\tau < \gamma'_\tau$; then $\gamma_t < \gamma'_t$ for all $t \geq \tau$. (2) In equilibrium corruption is always lower than in the static setting, $\gamma_t \leq \hat{\gamma}$, $\forall t$. (3) For all equilibrium sequences different from $\gamma_t = \gamma^*$, $\forall t$, $\beta_t^Y = \beta_t^{YC} < \beta_t^O < \beta^*$.

Proof. Appendix A.5.

Part 1 of Proposition 3 states that equilibrium sequences are monotonic in the sense that if we consider two different equilibrium sequences that have different values of corruption at a given date τ , both no lower than steady state corruption, γ^* , the ranking between levels of corruption is preserved at all subsequent dates. Part 2 is straightforward. Part 3 shows that outside of the steady state old agents have a higher propensity to become corrupt as $\beta_t^O > \beta_t^{YC}$ for all t .¹⁶ It should be stressed that this is not to say that old agents account for more than half of corruption at any point in time as the distribution of old agents has a lower mass of agents with low β , as some of them were detected as corrupt when young and replaced.

5 The dynamic response of corruption

One of the main reasons to develop a dynamic model of corruption is to be able to make predictions about the response of corruption to shocks that affect its profitability in temporary or permanent ways and as a consequence to establish the desirability of these shocks.

The purpose of this section is to use the results summarized in Propositions 2 and 3 to study equilibrium sequences starting after a temporary shock or after a permanent and unforeseen change of fundamentals that implies a different steady state level of corruption.

5.1 Response to temporary shocks

The following proposition characterizes the dynamics of corruption following a temporary shock to the profitability of corruption.¹⁷

Proposition 4 Suppose the system was in the steady state γ^* at $t = 0$ and that a temporary shock to some parameter (other than μ) occurs for period $t = 1$ only changing the profitability of corruption to $\tilde{\Gamma}(\gamma)$, (and becoming known before players play). Then:

¹⁶The same result has been obtained in a different model by Lui (1986) and is consistent with Rose-Ackerman's (1978) claim that lame-duck politicians are especially susceptible to bribes.

¹⁷As for the case of Proposition 6 below, in the following Proposition we ignore the cases in which the change in the excess profitability of corruption is determined by a change in the monitoring probability unforeseen until the *beginning* of period τ . A similar result can be proven using a slightly more involved argument. Notice, however, that Proposition 5 below considers the case of a change in the monitoring probability unforeseen until the *end* of period τ .

1. If the shock increased the profitability of corruption, $\tilde{\Gamma}(\gamma) > \Gamma(\gamma)$, $\forall \gamma \in [0, \hat{\gamma})$

$$\gamma_1 > \gamma_2 > \gamma_3 > \dots > \gamma^*.$$

2. If the shock decreased the profitability of corruption, $\tilde{\Gamma}(\gamma) < \Gamma(\gamma)$, $\forall \gamma \in [0, \hat{\gamma})$

$$\gamma_1 < \gamma^* \text{ and } \gamma_2 > \gamma_3 > \dots > \gamma^*.$$

Proof. Appendix A.6.

Part 1 of Proposition 4 shows that an increase in the profitability of corruption, even if it is temporary, may lead to increased levels of corruption over a long period of time, as the initial level of corruption above the steady state implies that the distribution of old agents' corruption costs in the following period satisfies condition of case (C1). As was argued in subsection 4.1, this implies that corruption in periods $t = 2, 3, \dots$ will be above the steady state, although converging to it.

An example of the case considered in part 1 of Proposition 4 is the wave of privatizations of state enterprises in many former communist bloc countries as well as India and Mexico that has created very profitable although temporary corruption opportunities.¹⁸ Proposition 4 shows that the negative effects in terms of corruption may well last beyond the privatization period and provides a rationale for the sustained levels of corruption that are being documented in those countries.¹⁹

Part 2 of Proposition 4, on the other hand, shows that a temporary decrease in the profitability of corruption, if beneficial in the short run, is counterproductive in the long run. This is a direct consequence of the fact that although a temporary decrease in the profitability of corruption leads to an instantaneous decrease of corruption, the fact that some agents that would be corrupt in the steady state choose in fact not to be corrupt in the period in which the shock takes place implies that the distribution of old agents' corruption costs in the following period will satisfy condition (C1) discussed on page 15 and corruption in all following periods will therefore be above the steady state although converging back down to it. It is even possible that the long run effects more than compensate for the short run gains in the sense that the average level of corruption following a temporary decrease in the profitability of corruption is higher than the steady state level of corruption.

While Proposition 4 deals with changes that were unforeseen until the beginning of a given period (so that they were unforeseen by the young when they made their choices in the previous period but not by the agents, young or old, who are currently alive), the next Proposition considers the case of a temporary

¹⁸Celarier (1997) provides interesting accounts of such corruption opportunities and how they have been exploited.

¹⁹The reader might find intriguing to know that part 1 of Proposition 4 also gives an accurate description of the dynamics of corruption in Hadleyburg following an unforeseen temporary corruption opportunity. After the sudden increase in corruption related in the tale, Hadleyburg eventually reverted to being "[...] an honest town once more. And the man will have to rise early to catch it napping again." Twain (1899), page 60.

change in the monitoring probability for one period that takes place at the end of the period, i.e., after agents currently alive have made their choices.²⁰ We consider this case as we are interested in the long run effects of short lived anti-corruption campaigns, such as an amnesty for firms that denounce corrupt agents (which implies a monitoring probability equal to 1), and of short spells of relaxed monitoring.

Proposition 5 *Suppose the system was in the steady state γ^* at $t = 0$ and that (unforeseen until the end of period 1) the probability of detection (for period 1 only) changes from μ to μ' .*

1. *If $\mu' > \mu$, then $\gamma_1 = \gamma^*$, $\gamma_2 < \gamma^*$ and $\gamma_3 > \gamma_4 > \dots > \gamma^*$;*
2. *If $\mu' < \mu$, then $\gamma_1 = \gamma^*$, $\gamma_2 > \gamma_3 > \gamma_4 > \dots > \gamma^*$.*

Proof. Appendix A.7.

Proposition 5 shows that, while the fact that the change in the monitoring probability μ is unforeseen until the end of period 1 implies that agents' decisions do not change (and therefore $\gamma_1 = \gamma^*$), the change in μ induces a change in the distribution of old agents' corruption costs in the following period. If $\mu' > \mu$, this distribution will satisfy condition (C2), and if $\mu' < \mu$, it will satisfy condition (C1) and the stated results on the continuation sequences are straightforward applications of Proposition 2.

Part 1 of Proposition 5 shows that short sweeping anticorruption campaigns, such as the Italian “Clean Hands” of 1992-93, are not only useless in the long run, they are counterproductive. With respect to this result it is interesting to notice that the recent political debate on the ultimate consequences of “Clean Hands” has voiced the concern that, as Part 1 of Proposition 5 predicts, apart from a short period of beneficial effects, Italian corruption is now even higher than before the anticorruption campaign.

Part 2 of Proposition 5 shows that even a short-lived decrease in the detection probability of corrupt agreements has long lasting consequences.

The previous results cast a doubt on the desirability of temporary shocks to policy parameters, such as μ , B_A , and B_F , as, while the immediate consequences may be positive or negative, any temporary deviation from the steady state has negative long run consequences in the sense that the level of corruption from the second period on will be higher than the steady state, even if converging to it.²¹

²⁰It is easy to see that any temporary unforeseen change other than in the monitoring probability that takes place after agents have decided to be corrupt or not can only have an impact on their payoff but not on the equilibrium.

²¹Tirole (1996) considers amnesties in a period in which a temporary shock to the excess profitability of corruption implies that all corruptible agents choose to be corrupt and concludes that they can be beneficial as they can allow society to stay in the low corruption equilibrium. While for Tirole (1996) an amnesty in one period is simply a way to make current decision unobservable, so that in the continuation equilibrium the economy stays in the steady state, in our model we can think of amnesties that either imply $\mu = 0$ (in the case in which corrupt public officials are pardoned) or $\mu = 1$ (in the case in which firms are pardoned if they give in their corruption partners) and the immediate consequences of these two different types of amnesties are shown to be very different.

5.2 Response to permanent shocks

The goal of this section is to study the dynamics of corruption following permanent changes in the profitability of corruption.

Our main result is stated in the following Proposition.

Proposition 6 *Suppose the system was in the steady state γ' at $t = 0$ and that a permanent shock to some parameter (other than μ) taking place in $t = 1$ (and becoming known before players play) implies a new steady state $\gamma^* \neq \gamma'$. Then*

1. *If $\gamma' < \gamma^*$ the equilibrium sequence $\{\gamma'_t\}_{t=1}^\infty$ will be such that:*

$$(a) \quad \widehat{\gamma} > \gamma'_1 > \gamma'_2 > \gamma'_3 > \dots > \gamma^*;$$

$$(b) \quad \gamma'_1 \text{ is decreasing in } \gamma'.$$

2. *If $\gamma' > \gamma^*$ the equilibrium sequence $\{\gamma'_t\}_{t=1}^\infty$ will be such that:*

$$(a) \quad \widehat{\gamma} > \gamma'_1 > \gamma'_2 > \gamma'_3 > \dots > \gamma^*;$$

$$(b) \quad \gamma'_1 \text{ is increasing in } \gamma'.$$

Proof. Appendix A.8.

Although Proposition 6 considers permanent changes in parameters other than the detection probability μ , Proposition 7 in Appendix A.9 shows that similar results can be obtained also for permanent changes in μ . Given changes in μ imply slight differences and significantly more involved formal arguments, the presentation of this result and its proof is relegated to the Appendix.

The intuition behind Proposition 6 is best explained referring separately to changes leading to higher and lower steady state levels of corruption and resorting to the intuition of Proposition 2.

When a shock implies that the new steady state is higher, the system inherits a distribution of old agents such that some agents that would be corrupt in the new steady state chose in fact not to be corrupt in the previous period. This implies that, as compared with the new steady state, some agents with a corruption cost below the new steady state indifference threshold are not replaced by agents that with positive probability would have a corruption cost higher than the new steady state indifference threshold. Since this implies that the old agents' distribution has a mass of agents with corruption costs less than or equal to the steady state indifference level that is higher than the corresponding mass in the steady state distribution, corruption will be higher than in the steady state. By Proposition 2 it is then easy to see that corruption will be decreasing towards the steady state. Since it can be seen that corruption will be higher,

the larger the departure of the inherited distribution from the distribution associated to the new steady state, it is easy to see that the lower the old steady state the higher the initial point of the equilibrium sequence as claimed in part 1b.

When a shock implies that the new steady state is lower, on the other hand, the system inherits a distribution of old agents such that some agents that would not be corrupt in the new steady state chose in fact to be corrupt in the previous period. This implies that, as compared with the new steady state, some agents with a corruption cost above the new steady state indifference threshold are replaced by agents that with positive probability will have a corruption cost lower than the new steady state indifference threshold. Since this implies that the old agents' distribution has a mass of agents with corruption costs less than or equal to the steady state indifference level that is higher than the corresponding mass in the steady state distribution, corruption will be higher than in the steady state. As before, by Proposition 2 corruption will be decreasing towards the steady state. Since, as argued above, corruption will be higher, the larger the departure of the inherited distribution from the distribution associated to the new steady state, the higher the old steady state the higher the initial point of the equilibrium sequence as claimed in part 2b.

Proposition 6 shows two important features of the adjustment to a new steady state following a permanent shock to the profitability of corruption:

1. The adjustment to the steady state can only occur through decreasing sequences and the level of corruption ensuing immediately after the shock has therefore to be above the new steady state.
2. There is an inherent asymmetry in the adjustment from an old steady state to a new one
 - (a) If the old steady state was lower than the new one, the new level of corruption (from which convergence to the new steady state starts) will be *higher*, the *lower* the old steady state.
 - (b) If the old steady state was higher than the new one, the new level of corruption (from which convergence to the new steady state starts) will be *higher*, the *higher* the old steady state.

An interesting case is the transition from a situation in which corruption is organized to a system of competitive corruption. Referring to the spectacular surge in corruption in former Soviet countries, Bardhan (1997) argues that “[...] the Communist Party used to centralize the collection of bribes and effectively monitored (sometimes with the help of the KGB) deviations from agreed-upon patterns of corruption. Now different ministries, agencies, and levels of local governments all set their own bribes independently in a decentralized attempt to maximize their own revenue.”²²

²²Bardhan (1997), page 1325.

A straightforward extension of our model (of competitive corruption) shows that the steady state level of corruption is lower and corruption revenue is higher under organized corruption, that is to say in a situation in which agents collude to decide the level of corruption, being aware that high levels of corruption imply a response of the affected party that decreases its profitability.²³ Proposition 6 therefore suggests that the immediate consequence of a transition from organized to competitive corruption in former Soviet countries may be a remarkable increase in the level of corruption over and above the new (already higher) steady state, although corruption should then decrease over time and eventually converge to the new steady state.

6 Conclusions

The corruption model we consider relies on three basic assumptions:

- A.1' Detection of corrupt activities leads to the replacement of corrupt agents and implies that the historical levels of corruption affect the distribution of the agents' cost to arrange corrupt transactions.
- A.2' The payoff from getting corrupt is decreasing in corruption.
- A.3' Young agents care about the future and therefore trade off current corruption payoff against the possible loss of future corruption opportunities resulting from discharge.

Although our results do not depend on the fine details of the model, it is important to highlight to what extent they depend on assumptions A.1', A.2', and A.3' above.

Assumption A.1' is a key element in that it is necessary to establish an intertemporal link between past and current corruption. Different assumptions establishing an intertemporal link lead to similar results. For instance, increasing penalty schedules, with stronger penalties associated with reiterated wrong-doing have been defended at least since Stigler's (1970) seminal work as providing marginal deterrence and as being justified in the light of imperfect observability. In Celentani and Ganuza (1997) we consider a situation in which agents face this kind of penalty schedules and obtain results similar to the ones presented in this paper.

Assumption A.2' is also essential for the following reasons. It is essential that the payoff of corruption is decreasing in the (aggregate) level of corruption rather than constant in order to have long lasting consequences of shocks: If the payoff of corruption were constant, a shock could only have an impact in the period in which it occurs and in the subsequent one through its impact on the distribution of young agents. It is also essential that the payoff of corruption is decreasing in the (aggregate) level of corruption rather

²³ A similar prediction has also been done by Shleifer and Vishny (1993) although using a different argument.

than increasing as in the latter case there would be more than one steady state and our results would not hold.

As was suggested in the discussion of Proposition 2 in subsection 4.1, when agents are completely myopic, the results on the dynamics of the aggregate level of corruption are qualitatively the same. In this sense Assumption A.3', while realistic, is not essential in deriving most of our results. It is, however, reassuring to find out that allowing young agents to condition their current decisions on their expectations about the future does not imply qualitative changes in the characterization of the dynamic properties of equilibrium sequences and in particular in their uniqueness. Assumption A.3' moreover is useful in highlighting that the fact that young agents do care about the future has two important implications: (1) There is a difference in the corruption decisions of young and old agents; in particular we show that while old agents are more prone to corruption, in the sense that their indifference threshold is higher than the one of young agents for a given level of aggregate corruption, the replacement of corrupt agents implies that the density of old agents below the indifference threshold is lower. (2) More importantly, given equilibrium sequences of corruption are decreasing and the profitability of corruption is therefore increasing over time, if young agents care about future payoffs, some of them will refrain from becoming corrupt when young in anticipation of higher payoff when old. As was argued in subsection 4.1 this has a direct effect that tends to reduce the level of corruption, in much the same way in which steady state corruption is lower than corruption in a static framework, but also generates a more unfavorable future distribution of old agents and has, therefore, an indirect effect that tends to increase corruption and ultimately leads to a reduction in the speed of convergence to the steady state.

Our work shows that under reasonable assumptions, the dynamics of corruption following exogenous shocks can be different from the intuition that could be derived by simple static models. An interesting direction for future research is to analyze the effect on equilibrium corruption of anticipated volatility of policy regimes defined by penalties and monitoring probabilities. We conjecture that the Hadleyburg effect implies the undesirability of the volatility of policy regimes, as such volatility would induce average levels of corruption higher than the ones that would obtain under a deterministic policy regime with equal expected penalties.

Appendices

A.1 Computation of $\beta_t^Y(q_t, q_{t+1})$

Let β denote a type of young agent in t who plans to be corrupt in $t + 1$ and who is contemplating getting corrupt in t . In order to determine whether such a type will be corrupt in t or not we first need to compute his expected payoff in $t + 1$ if he is not fired. It is easy to show that

$$\pi_A(q_{t+1}) = b_{t+1} - \mu B_A - \beta. \quad (27)$$

From the generalized Nash bargaining solution we get

$$b_{t+1} = \alpha[C(q_{t+1}) - C(q_c) - \mu B_F] + (1 - \alpha)[\beta + \mu B_A] \quad (28)$$

and substituting (28) into (27) we get

$$\pi_A(q_{t+1}) = \alpha[C(q_{t+1}) - k - \beta]$$

Given this, the bargaining situation confronting at date t a young agent with the firm when the quality standard is q_t can be described by the following generalized Nash bargaining problem in which both the payoff from getting corrupt in t and the disagreement point for the agent now include the expected payoff from corruption in $t + 1$:

$$\max_{b_t} (b_t - \mu B_A + (1 - \mu) \max\{\pi_A(q_{t+1}), 0\} - \max\{\pi_A(q_{t+1}), 0\})^\alpha (C(q_{t+1}) - b_t - \mu B_F - C(q_c))^{1-\alpha})$$

Straightforward although tedious calculations show that, given q_t and q_{t+1} , the bribe in t will be

$$b_t = \alpha(C(q_t) - \mu B_F - C(q_c)) + (1 - \alpha)(\beta + \mu B_A + \mu \max\{\pi_A(q_{t+1}), 0\})$$

Now let β_t^{YC} denote the type of young agent at date t that plans to be corrupt at $t + 1$ ²⁴ and who is exactly indifferent between being corrupt or not today. To find such type we need to equate the payoffs from being corrupt and honest in t when the agent anticipates that, if not caught, he will be corrupt in $t + 1$

$$b_t - \beta_t^{YC} - \mu B_A + (1 - \mu) \pi_A(q_{t+1}) = \pi_A(q_{t+1}). \quad (29)$$

Straightforward calculations show that equation (29) implies that

$$\beta_t^{YC} = \frac{C(q_t) - \mu \alpha C(q_{t+1})}{1 - \mu \alpha} - k,$$

As before we assume without loss of generality that $\beta_t^{YC} \leq 1$. Since by definition, $\beta_t^{YC} < \beta_{t+1}^O$, we define

²⁴For such a type of agent $\max\{\pi_A(q_{t+1}), 0\} = \pi_A(q_{t+1})$.

$$\beta_t^{YC} = \begin{cases} \frac{C(q_t) - \mu\alpha C(q_{t+1})}{1 - \mu\alpha} - k & \text{if } \frac{C(q_t) - \mu\alpha C(q_{t+1})}{1 - \mu\alpha} - k < \beta_{t+1}^O \\ 0 & \text{otherwise} \end{cases}$$

where $\beta_t^{YC} = 0$ means that, given q_t and q_{t+1} , there is no type of agent who is indifferent between being corrupt or not in t and who plans to be corrupt in $t+1$. The above definition can be shown to be equivalent to

$$\beta_t^{YC} = \begin{cases} \frac{C(q_t) - \mu\alpha C(q_{t+1})}{1 - \mu\alpha} - k & \text{if } q_t < q_{t+1} \\ 0 & \text{otherwise.} \end{cases} \quad (30)$$

If a young agent in t plans not to be corrupt in $t+1$, his behavior will be identical to that of an old agent in t (recall that we have normalized wage to 0). Let β_t^{YH} denote a type of young agent in t who plans to be honest in $t+1$.²⁵ Noticing that by definition $\beta_t^{YH} > \beta_{t+1}^O$ we can then define

$$\beta_t^{YH} = \begin{cases} C(q_t) - k & \text{if } C(q_t) - k > \beta_{t+1}^O \\ 0 & \text{otherwise} \end{cases}$$

where $\beta_t^{YH} = 0$ means that, given q_t and q_{t+1} , there is no type of agent who is indifferent between being corrupt or not in t and who plans to be honest in $t+1$. The above definition can be shown to be equivalent to

$$\beta_t^{YH} = \begin{cases} 0 & \text{if } q_t < q_{t+1} \\ C(q_t) - k & \text{if } q_t \geq q_{t+1} \end{cases}. \quad (31)$$

As always we assume without loss of generality that $\beta_t^{YH} \leq 1$.

Letting $\beta_t^Y(q_t, q_{t+1}) = \max\{\beta_t^{YH}, \beta_t^{YC}\}$ gives equation (10).

A.2 Derivation of equation (16)

Since the fraction of corrupt agents in each period is equal to the average of corrupt old and young agents, to compute it starting from equations (15) and (11) we need to consider four different cases that depend on how quality standards and therefore corruption change over time.

If $q_{t-1} > q_t$ and $q_t < q_{t+1}$, average corruption in t , γ_t , will be:

$$\begin{aligned} \gamma_t &= \frac{1}{2} (\gamma_t^Y + \gamma_t^O) \\ &= \frac{1}{2} [\beta_t^Y + (1 - \mu) \beta_t^O + \mu \beta_{t-1}^Y \beta_t^O] \\ &= \frac{1}{2} \left[(2 - \mu)(C(q_t) - k) + \mu(C(q_{t-1}) - k)(C(q_t) - k) - \frac{\mu\alpha}{1 - \mu\alpha} (C(q_{t+1}) - C(q_t)) \right] \end{aligned}$$

If $q_{t-1} < q_t < q_{t+1}$, then:

$$\gamma_t = \frac{1}{2} (\gamma_t^Y + \gamma_t^O)$$

²⁵ For such a type of agent $\max\{\pi_A(q_{t+1}), 0\} = 0$.

$$\begin{aligned}
&= \frac{1}{2} \left[\beta_t^Y + \beta_t^O - \mu \beta_{t-1}^Y + \mu \beta_{t-1}^Y \beta_t^O \right] \\
&= \frac{1}{2} \left[(2 - \mu)(C(q_t) - k) + \mu(C(q_{t-1}) - k)(C(q_t) - k) - \frac{\mu\alpha}{1 - \mu\alpha} (C(q_{t+1}) - C(q_t)) \right. \\
&\quad \left. - \frac{\mu}{1 - \mu\alpha} (C(q_{t-1}) - C(q_t))(1 - \mu\alpha(C(q_t) - k)) \right]
\end{aligned}$$

If $q_{t-1} > q_t > q_{t+1}$, then:

$$\begin{aligned}
\gamma_t &= \frac{1}{2} (\gamma_t^Y + \gamma_t^O) \\
&= \frac{1}{2} \left[\beta_t^Y + (1 - \mu) \beta_t^O + \mu \beta_{t-1}^Y \beta_t^O \right] \\
&= \frac{1}{2} \left[(2 - \mu)(C(q_t) - k) + \mu(C(q_{t-1}) - k)(C(q_t) - k) \right]
\end{aligned}$$

Finally if $q_{t-1} < q_t$ and $q_t > q_{t+1}$, we have:

$$\begin{aligned}
\gamma_t &= \frac{1}{2} (\gamma_t^Y + \gamma_t^O) \\
&= \frac{1}{2} \left[\beta_t^Y + \beta_t^O - \mu \beta_{t-1}^Y + \mu \beta_{t-1}^Y \beta_t^O \right] \\
&= \frac{1}{2} \left[(2 - \mu)(C(q_t) - k) + \mu(C(q_{t-1}) - k)(C(q_t) - k) - \frac{\mu}{1 - \mu\alpha} (C(q_{t-1}) - C(q_t))(1 - \mu\alpha(C(q_t) - k)) \right].
\end{aligned}$$

It is easy to see that equation (16) summarizes the previous equations.

A.3 Proof of Proposition 1

(1) Let $e(\gamma) = (2 - \mu)\Gamma(\gamma) + \mu[\Gamma(\gamma)]^2 - 2\gamma$. The steady state is then given by the solution of $e(\gamma^*) = 0$ for $\gamma^* \in [0, 1]$. Since we assume that $\Gamma(0) > 0$ (otherwise no agent would ever want to be corrupt) to show part 1 of the Proposition it suffices to show that $e(0) > 0$, $e(\bar{\gamma}) < 0$, where $\bar{\gamma} = \min\{1, \tilde{\gamma}\}$ and $\tilde{\gamma} = \max\{\gamma : \Gamma(\gamma) \geq 0\}$ and $e'(\gamma) < 0$. It is easy to see that $e(0) = (2 - \mu)\Gamma(0) + \mu(\Gamma(0))^2 > 0$; On the other hand

$$e(\bar{\gamma}) = (2 - \mu)\Gamma(\bar{\gamma}) + \mu(\Gamma(\bar{\gamma}))^2 - 2\bar{\gamma}$$

If $\bar{\gamma} = 1$, we have

$$e(\bar{\gamma}) = (2 + \mu\Gamma(1))(\Gamma(1) - 1) < 0$$

with the inequality following from $0 \leq \Gamma(\bar{\gamma}) \leq \Gamma(1) < 1$; If $\bar{\gamma} = \tilde{\gamma}$, we have

$$e(\bar{\gamma}) = -2\tilde{\gamma} < 0.$$

Noticing that $e'(\gamma) = (2 - \mu)\Gamma'(\gamma) + 2\mu\Gamma(\gamma)\Gamma'(\gamma) - 2 < 0$ concludes the argument.

(2) Omitted.

(3) To explicitly take into account the impact of the parameters of the problem on the steady state, rewrite the equation of the steady state as $e(\gamma^*, \mu, B_A, B_F, q_c) = 0$. Recall that $\Gamma(\gamma) = C(q(\gamma)) - k$

and that $k = C(q_c) + \mu(B_A + B_F)$ which implies that $\Gamma(\gamma) = C(q(\gamma)) - C(q_c) - \mu(B_A + B_F)$. Let $a \in \{B_A, B_F, q_c\}$; then we have

$$\frac{d\gamma^*}{da} = -\frac{\frac{\partial e(\gamma^*)}{\partial a}}{\frac{\partial e(\gamma^*)}{\partial \gamma}} = -\frac{\frac{\partial e(\gamma^*)}{\partial \Gamma} \frac{\partial \Gamma(\gamma^*)}{\partial a}}{\frac{\partial e(\gamma^*)}{\partial \gamma}}.$$

Since $\frac{\partial e(\gamma^*)}{\partial \gamma} < 0$, $\frac{\partial e(\gamma^*)}{\partial \Gamma} > 0$, $\frac{\partial \Gamma(\gamma^*)}{\partial a} < 0$, for all $a \in \{B_A, B_F, q_c\}$ we have that $\frac{d\gamma^*}{da} < 0$ for all $a \in \{B_A, B_F, q_c\}$. Similarly, we have

$$\frac{d\gamma^*}{d\mu} = -\frac{\frac{\partial e(\gamma^*)}{\partial \mu}}{\frac{\partial e(\gamma^*)}{\partial \gamma}} = -\frac{\frac{\partial e(\gamma^*)}{\partial \Gamma} \frac{\partial \Gamma(\gamma^*)}{\partial \mu} + \frac{\partial e(\gamma^*)}{\partial \mu}}{\frac{\partial e(\gamma^*)}{\partial \gamma}} = -\frac{\frac{\partial e(\gamma^*)}{\partial \Gamma} \frac{\partial \Gamma(\gamma^*)}{\partial \mu} - [\Gamma(\gamma^*) - (\Gamma(\gamma^*))^2]}{\frac{\partial e(\gamma^*)}{\partial \gamma}} < 0$$

as $\frac{\partial \Gamma(\gamma^*)}{\partial \mu} < 0$ and $\Gamma(\gamma^*) - (\Gamma(\gamma^*))^2 > 0$ imply that the numerator is negative. ■

A.4 Proof of Proposition 2

Let

$$E(\gamma_t, \gamma_{t-1}) = (2 - \mu)\Gamma(\gamma_t) + \mu\Gamma(\gamma_{t-1})\Gamma(\gamma_t) - 2\gamma_t \quad (32)$$

$$E_P(\gamma_t, \gamma_{t-1}) = -\frac{\mu}{1 - \mu\alpha} [\Gamma(\gamma_{t-1}) - \Gamma(\gamma_t)] (1 - \mu\alpha\Gamma(\gamma_t)) \quad (33)$$

$$E_F(\gamma_t, \gamma_{t+1}) = -\frac{\mu\alpha}{1 - \mu\alpha} [\Gamma(\gamma_{t+1}) - \Gamma(\gamma_t)] \quad (34)$$

we can rewrite equation (17) as

$$E(\gamma_t, \gamma_{t-1}) + \delta_F E_F(\gamma_t, \gamma_{t+1}) + \delta_P E_P(\gamma_t, \gamma_{t-1}) = 0 \quad (35)$$

where subindices refer as before to the future and past dynamics of corruption: $\delta_F = 1$ if $\gamma_t > \gamma_{t+1}$, and $\delta_F = 0$ otherwise; $\delta_P = 1$ if $\gamma_{t-1} > \gamma_t$, and $\delta_P = 0$ otherwise. From (32) we get

$$\frac{\partial E(\gamma_t, \gamma_{t-1})}{\partial \gamma_t} = [(2 - \mu) + \mu\Gamma(\gamma_{t-1})] \Gamma'(\gamma_t) - 2 < 0 \quad (36)$$

$$\frac{\partial E(\gamma_t, \gamma_{t-1})}{\partial \gamma_{t-1}} = \mu\Gamma'(\gamma_{t-1})\Gamma(\gamma_t) < 0. \quad (37)$$

From (33) we get

$$\frac{\partial E_P(\gamma_t, \gamma_{t-1})}{\partial \gamma_{t-1}} = -\frac{\mu}{1 - \mu\alpha} \Gamma'(\gamma_{t-1}) (1 - \mu\alpha\Gamma(\gamma_t)) > 0. \quad (38)$$

From (34) we get

$$\frac{\partial E_F(\gamma_t, \gamma_{t+1})}{\partial \gamma_t} = \frac{\mu\alpha}{1 - \mu\alpha} \Gamma'(\gamma_t) < 0 \quad (39)$$

$$\frac{\partial E_F(\gamma_t, \gamma_{t+1})}{\partial \gamma_{t+1}} = -\frac{\mu\alpha}{1 - \mu\alpha} \Gamma'(\gamma_{t+1}) > 0 \quad (40)$$

Finally notice that

$$\frac{\partial [E(\gamma_t, \gamma_{t-1}) + E_P(\gamma_t, \gamma_{t-1})]}{\partial \gamma_{t-1}} = \mu\Gamma'(\gamma_{t-1}) \left[\Gamma(\gamma_t) - \frac{1 - \mu\alpha\Gamma(\gamma_t)}{1 - \mu\alpha} \right] > 0 \quad (41)$$

since $\Gamma(\gamma_t) < 1$ and $\frac{1-\mu\alpha\Gamma(\gamma_t)}{1-\mu\alpha} > 1$.

Given these preliminaries we will now show the result through a sequence of Lemmas. We will first show that no strictly nonmonotonic sequence can arise in equilibrium (Lemmas 1 and 3) and that no strictly increasing sequence can arise in equilibrium either (Lemma 2). We will then show that in an equilibrium sequence we cannot have $(\gamma_{t-1}, \gamma_t, \gamma_{t+1})$ with two adjacent levels of corruption equal to each other and the remaining different from them (Lemma 4). Having shown this, we will conclude that equilibrium sequences different from infinite repetitions of the steady state have to be strictly decreasing for all $t \geq 1$ (Lemma 5) which implies that equilibrium sequences have to be above the steady state.

Relying on the fact that equilibrium sequences different from infinite repetitions of the steady state are strictly decreasing for all $t \geq 1$ and bounded below by the steady state, we will show in Lemma 6 that the level of corruption at date $t = 0$ will be equal to, higher or lower than the steady state, depending on the initial distribution $G(\beta)$ and in particular on which of conditions (19), (20), or (21) holds.

We omit the proof of existence of equilibrium sequences which is a straightforward implication of the assumed continuity of the initial distribution $G(\beta)$ and of $F(\beta; \beta', \mu)$ for all β' and for all μ . Lemma 7 concludes the proof of Proposition 2 by showing that equilibrium sequences are unique.

Lemma 1 *Consider an equilibrium sequence $\{\gamma_t\}$. Then there is no $t \geq 1$ such that*

$$\gamma_{t-1} > \gamma_t \tag{42}$$

$$\gamma_t < \gamma_{t+1}. \tag{43}$$

Proof. Suppose that there is a $t \geq 1$ such that (42) and (43) hold. This implies that $\delta_F = 0$ and $\delta_P = 1$, and from equation (35) we have

$$E(\gamma_t, \gamma_{t-1}) + E_P(\gamma_t, \gamma_{t-1}) = 0. \tag{44}$$

Since $\gamma_{t-1} > \gamma_t$, $E_P(\gamma_t, \gamma_{t-1}) > 0$ and, in turn, $E(\gamma_t, \gamma_{t-1}) < 0$ which, together with (36) and (37) implies that $\gamma_{t-1} > \gamma^*$. Consider the following two cases:

Case 1.a: $\gamma_{t-1} > \gamma^*$, $\gamma_t < \gamma^*$. From above $E_P(\gamma_t, \gamma^*) > 0$ when $\gamma_t < \gamma^*$; then notice that when $\gamma_t < \gamma^*$, $E(\gamma^*, \gamma^*) = 0$ and (36) imply $E(\gamma_t, \gamma^*) > 0$. This implies that for $\gamma_t < \gamma^*$, $E(\gamma_t, \gamma^*) + E_P(\gamma_t, \gamma^*) > 0$ and, given (41), we have that $E(\gamma_t, \gamma_{t-1}) + E_P(\gamma_t, \gamma_{t-1}) > 0$ for $\gamma_{t-1} > \gamma^*$, $\gamma_t < \gamma^*$ thus contradicting (44).

Case 1.b: $\gamma_{t-1} > \gamma^*$, $\gamma_t > \gamma^*$. Consider the following two subcases

1. $\gamma_{t+2} > \gamma_{t+1}$. In this case from equation (35) we have

$$E(\gamma_{t+1}, \gamma_t) = 0$$

which is impossible since $E(\gamma_{t+1}, \gamma_t) < 0$ for $\gamma_{t+1} > \gamma_t > \gamma^*$.

2. $\gamma_{t+2} < \gamma_{t+1}$. In this case from equation (35) we have

$$E(\gamma_{t+1}, \gamma_t) + E_F(\gamma_{t+1}, \gamma_{t+2}) = 0; \quad (45)$$

Since $\gamma_{t+2} < \gamma_{t+1}$ implies $E_F(\gamma_{t+1}, \gamma_{t+2}) < 0$, for (45) to hold we need $E(\gamma_{t+1}, \gamma_t) > 0$ which is impossible since $\gamma_{t+1} > \gamma_t > \gamma^*$, (36) and (37) imply $E(\gamma_{t+1}, \gamma_t) < 0$. ■

Lemma 2 Consider an equilibrium sequence $\{\gamma_t\}$. Then there is no $t \geq 1$ such that

$$\gamma_{t-1} < \gamma_t \quad (46)$$

$$\gamma_t < \gamma_{t+1} \quad (47)$$

Proof. Suppose that there is a $t \geq 1$ such that (46) and (47) hold. This implies that $\delta_F = 0$ and $\delta_P = 0$, and from equation (35) we have

$$E(\gamma_t, \gamma_{t-1}) = 0.$$

This implies that $\gamma_{t-1} < \gamma^*$ and $\gamma_t > \gamma^*$ and the latter together with (47) implies $\gamma_{t+1} > \gamma^*$. Consider the following two cases.

Case 2.a: $\gamma_{t+2} \geq \gamma_{t+1}$. In this case from equation (35) we have

$$E(\gamma_{t+1}, \gamma_t) = 0$$

which is impossible since for $\gamma_{t+1} > \gamma_t > \gamma^*$, $E(\gamma_{t+1}, \gamma_t) < 0$.

Case 2.b: $\gamma_{t+2} < \gamma_{t+1}$. In this case from equation (35) we have

$$E(\gamma_{t+1}, \gamma_t) + E_F(\gamma_{t+1}, \gamma_{t+2}) = 0 \quad (48)$$

Since $\gamma_{t+2} < \gamma_{t+1}$, $E_F(\gamma_{t+1}, \gamma_{t+2})$ for (48) to hold we need, $E(\gamma_{t+1}, \gamma_t) > 0$ which is impossible since for $\gamma_{t+1} > \gamma_t > \gamma^*$, $E(\gamma_{t+1}, \gamma_t) < 0$. ■

Lemma 3 Consider an equilibrium sequence $\{\gamma_t\}$. Then there is no $t \geq 2$ such that

$$\gamma_{t-1} < \gamma_t \quad (49)$$

$$\gamma_t > \gamma_{t+1} \quad (50)$$

Proof. Suppose that there is a $t \geq 2$ such that (49) and (50) hold, and consider the following three subcases:

1. $\gamma_{t-2} < \gamma_{t-1}$. Since $\gamma_{t-1} < \gamma_t$ we are in the case of Lemma 2 which rules out such a subsequence.

2. $\gamma_{t-2} > \gamma_{t-1}$. Since $\gamma_{t-1} < \gamma_t$ we are in the case of Lemma 1 which rules out such a subsequence.
3. $\gamma_{t-2} = \gamma_{t-1}$. Since $\gamma_{t-1} < \gamma_t$, this case is ruled out by Lemma 4 below. ■

Lemma 4 *Let $\{\gamma_t\}$ denote an equilibrium sequence. Then there is no $t \geq 1$ such that $\gamma_{t-1} = \gamma_t \neq \gamma_{t+1}$ or such that $\gamma_{t-1} \neq \gamma_t = \gamma_{t+1}$.*

Proof. We will only show that there is no $t \geq 1$ such that $\gamma_{t-1} = \gamma_t \neq \gamma_{t+1}$ arises in equilibrium as the proof of the second part of the statement follows along the same lines. Suppose that $\gamma_{t-1} = \gamma_t \neq \gamma_{t+1}$ and consider the following two cases:

1. $\gamma_{t-1} = \gamma_t < \gamma_{t+1}$

Then from (35) we have

$$E(\gamma_t, \gamma_{t-1}) = 0$$

which is satisfied for $\gamma_{t-1} = \gamma_t$ only if $\gamma_{t-1} = \gamma_t = \gamma^*$. Now consider the following three subcases:

- (a) $\gamma^* = \gamma_t < \gamma_{t+1} < \gamma_{t+2}$. This case is ruled out by Lemma 2.
- (b) $\gamma^* = \gamma_t < \gamma_{t+1}, \gamma_{t+1} > \gamma_{t+2}$. In this case

$$E(\gamma_{t+1}, \gamma_t) + E_F(\gamma_{t+1}, \gamma_{t+2}) = 0$$

which is impossible as both terms are strictly negative.

- (c) $\gamma^* = \gamma_t < \gamma_{t+1} = \gamma_{t+2}$. From equation (35) we then have

$$E(\gamma_{t+1}, \gamma^*) = 0$$

which implies that $\gamma_{t+1} = \gamma^*$ thus violating the assumption that $\gamma_t < \gamma_{t+1}$.

2. $\gamma_{t-1} = \gamma_t > \gamma_{t+1}$

Then from (35) we have

$$E(\gamma_t, \gamma_{t-1}) + E_F(\gamma_t, \gamma_{t+1}) = 0.$$

Since $\gamma_t > \gamma_{t+1}$, $E_F(\gamma_t, \gamma_{t+1}) < 0$ which implies that $E(\gamma_t, \gamma_{t-1}) > 0$ which together with $E(\gamma^*, \gamma^*) = 0$ and (36) and (37) implies that $\gamma_t = \gamma_{t-1} < \gamma^*$. Now consider the following three subcases:

- (a) $\gamma^* = \gamma_t > \gamma_{t+1} > \gamma_{t+2}$. This case is ruled out by Lemma 5 below that shows that strictly decreasing sequences have to lie above the steady state..

(b) $\gamma^* = \gamma_t > \gamma_{t+1}$, $\gamma_{t+1} < \gamma_{t+2}$. This case is ruled out by Lemma 1.

(c) $\gamma^* = \gamma_t > \gamma_{t+1} = \gamma_{t+2}$. From equation (35) we then have

$$E(\gamma_{t+1}, \gamma_t) + E_P(\gamma_{t+1}, \gamma_t) = 0.$$

Since $\gamma_t > \gamma_{t+1}$, $E_P(\gamma_{t+1}, \gamma_t) > 0$ which implies that $E(\gamma_{t+1}, \gamma_t) < 0$. However $E(\gamma^*, \gamma^*) = 0$, $\gamma^* = \gamma_t > \gamma_{t+1}$, and (37) imply that $E(\gamma_{t+1}, \gamma_t) > 0$, thus leading to a contradiction. ■

The following Lemma shows that all equilibrium sequences have to be either constant or strictly decreasing for $t \geq 1$.

Lemma 5 *Equilibrium sequences $\{\gamma_t\}_{t=0}^\infty$ different from $\{\gamma'_t\}_{t=0}^\infty$, $\gamma'_t = \gamma^*$ for $t = 0, 1, \dots$, are such that, for all $t \geq 2$, $\gamma_t > \gamma^*$ and*

$$\gamma_{t-1} > \gamma_t \tag{51}$$

$$\gamma_t > \gamma_{t+1}. \tag{52}$$

Proof. Straightforward from Lemmas 1, 2.3 and 4. ■

Using the previous results the following Lemma characterizes the equilibrium level of corruption at $t = 0$ given an arbitrary initial distribution.

Lemma 6 *Let $G(\beta)$ be given and let $\{\gamma'_t\}_{t=0}^\infty$ denote an equilibrium sequence starting from this distribution.*

1. *If*

$$G(\beta^*) = F(\beta^*; \beta^*, \mu) \tag{53}$$

then $\{\gamma'_t\}_{t=0}^\infty$ with $\gamma'_t = \gamma^$, $t = 0, 1, \dots$*

2. *If*

$$G(\beta^*) > F(\beta^*; \beta^*, \mu) \tag{54}$$

then $\{\gamma'_t\}_{t=0}^\infty$ with $\gamma'_0 > \gamma'_1 > \gamma'_2 > \dots > \gamma^$.*

3. *If*

$$G(\beta^*) < F(\beta^*; \beta^*, \mu) \tag{55}$$

then $\{\gamma'_t\}_{t=0}^\infty$ with $\gamma'_0 < \gamma^$, $\gamma'_1 > \gamma'_2 > \dots > \gamma^*$.*

Proof. (1) Recalling that $\beta^* = \Gamma(\gamma^*)$, in the steady state we have

$$\gamma^* = \frac{1}{2} [\Gamma(\gamma^*) + F(\Gamma(\gamma^*); \beta^*, \mu)]. \tag{56}$$

Suppose now that $\{\gamma'_t\}_{t=0}^\infty$, with $\gamma'_t = \gamma^*$, $t = 0, 1, \dots$. This implies that in period $t = 0$, we have

$$\gamma^* = \frac{1}{2} [\Gamma(\gamma^*) + G(\Gamma(\gamma^*))]$$

which is clearly satisfied given (56) as, by assumption, $G(\Gamma(\gamma^*)) = F(\Gamma(\gamma^*); \beta^*, \mu)$.

(2) From Lemma 5 we know that $\gamma_1 > \gamma_2 > \dots > \gamma^*$. We therefore only need to show that if (54) holds, $\gamma_0 > \gamma^*$. To do this suppose (54) holds and, contrary to the statement, $\gamma_0 < \gamma^*$. Given this implies that $\gamma_1 > \gamma_0$, the profitability of corruption is higher at $t = 0$ than at $t = 1$, and this implies that all agents, both young and old, with types $\beta < \Gamma(\gamma_0)$ will get corrupt so that γ_0 will be determined by the following equation:

$$\gamma_0 = \frac{1}{2} [\Gamma(\gamma_0) + G(\Gamma(\gamma_0))]. \quad (57)$$

If $\gamma_0 < \gamma^*$

$$\Gamma(\gamma_0) > \Gamma(\gamma^*) \quad (58)$$

which together with (54) implies that

$$G(\Gamma(\gamma_0)) > F(\Gamma(\gamma^*); \beta^*, \mu). \quad (59)$$

It is now easy to see that (56), (58) and (59) provide a contradiction to (57) when $\gamma_0 < \gamma^*$.

(3) From Lemma 5 we know that $\gamma_1 > \gamma_2 > \dots > \gamma^*$. We therefore only need to show that if (55) holds, $\gamma_0 < \gamma^*$. Suppose, contrary to the statement that under (55) $\gamma_0 > \gamma^*$.

We have to consider two cases: (a) $\gamma_0 \leq \gamma_1$; (b) $\gamma_0 > \gamma_1$.

(a) A contradiction to the hypothesis that $\gamma_0 > \gamma^*$ is reached using the same arguments that were used in the proof of part 2, with all inequalities reversed.

(b) Since $\gamma_0 > \gamma_1 > \gamma^*$ the law of motion is

$$\gamma_0 = \frac{1}{2} [\Gamma(\gamma_0) + G(\Gamma(\gamma_0)) + E_F(\gamma_0, \gamma_1)]. \quad (60)$$

Since $\gamma_0 > \gamma_1$, $E_F(\gamma_0, \gamma_1) < 0$ which, together with (60), implies

$$\gamma_0 < \frac{1}{2} [\Gamma(\gamma_0) + G(\Gamma(\gamma_0))]. \quad (61)$$

Under the assumption that $\gamma_0 > \gamma^*$ we have

$$\Gamma(\gamma_0) < \Gamma(\gamma^*) \quad (62)$$

which, together with (55), implies

$$G(\Gamma(\gamma_0)) < F(\Gamma(\gamma^*); \beta^*, \mu). \quad (63)$$

and it is easy to see that (56), (62), and (63) provide a contradiction to (61) when $\gamma_0 > \gamma^*$. ■

The following Lemma shows that equilibrium sequences are unique and concludes the proof of Proposition 2.

Lemma 7 *Let $G(\beta)$ be given. Then there exists a unique equilibrium sequence $\{\gamma_t\}_{t=0}^\infty$ starting from this distribution.*

Proof. To prove the Lemma it is sufficient to show that given an arbitrary distribution $G(\beta)$ no two equilibrium sequences exist, $\{\gamma_t\}_{t=0}^\infty$ and $\{\gamma'_t\}_{t=0}^\infty$, with $\gamma_0 \neq \gamma'_0$. Consider a distribution $G(\beta)$ with $G(\beta^*) > F(\beta^*; \beta^*, \mu)$ and suppose that the previous statement is not true, i.e., suppose that, given $G(\beta)$, there are two equilibrium sequences $\{\gamma_t\}_{t=0}^\infty$ and $\{\gamma'_t\}_{t=0}^\infty$ with $\gamma_0 \neq \gamma'_0$. Suppose without loss of generality that $\gamma_0 > \gamma'_0$. Since we know from Lemma 5 that equilibrium sequences have to be strictly decreasing, the two following equations must hold

$$\gamma_0 = \frac{1}{2} [\Gamma(\gamma_0) + G(\Gamma(\gamma_0)) + E_F(\gamma_0, \gamma_1)] \quad (64)$$

$$\gamma'_0 = \frac{1}{2} [\Gamma(\gamma'_0) + G(\Gamma(\gamma'_0)) + E_F(\gamma'_0, \gamma'_1)]. \quad (65)$$

Recalling that

$$E_F(\gamma_t, \gamma_{t+1}) = -\frac{\mu\alpha}{1-\mu\alpha} [\Gamma(\gamma_{t+1}) - \Gamma(\gamma_t)] \quad (66)$$

it is easy to see that (64) and (65) are equivalent to

$$2\gamma_0 - \frac{1}{1-\mu\alpha} \Gamma(\gamma_0) + \frac{\mu\alpha}{1-\mu\alpha} \Gamma(\gamma_1) = G(\Gamma(\gamma_0)) \quad (67)$$

$$2\gamma'_0 - \frac{1}{1-\mu\alpha} \Gamma(\gamma'_0) + \frac{\mu\alpha}{1-\mu\alpha} \Gamma(\gamma'_1) = G(\Gamma(\gamma'_0)). \quad (68)$$

Since by assumption $\gamma_0 > \gamma'_0$ and since $\Gamma(\gamma)$ is decreasing and $G(\beta)$ we have $\Gamma(\gamma_0) < \Gamma(\gamma'_0)$ and

$$G(\Gamma(\gamma_0)) < G(\Gamma(\gamma'_0)). \quad (69)$$

From (67), (68) and (69) we have

$$2\gamma'_0 - \frac{1}{1-\mu\alpha} \Gamma(\gamma'_0) + \frac{\mu\alpha}{1-\mu\alpha} \Gamma(\gamma'_1) > 2\gamma_0 - \frac{1}{1-\mu\alpha} \Gamma(\gamma_0) + \frac{\mu\alpha}{1-\mu\alpha} \Gamma(\gamma_1) \quad (70)$$

which is equivalent to

$$\Gamma(\gamma'_1) - \Gamma(\gamma_1) > 2\frac{1-\mu\alpha}{\mu\alpha} (\gamma_0 - \gamma'_0) + \frac{1}{\mu\alpha} (\Gamma(\gamma'_0) - \Gamma(\gamma_0)). \quad (71)$$

Since both terms on the right hand side of inequality (71) are positive, and recalling that $\Gamma(\gamma)$ is decreasing, we have that

$$\gamma_1 > \gamma'_1. \quad (72)$$

Consider now corruption in period $t = 1$. Since all young agents with corruption cost below $\frac{\Gamma(\gamma_0) - \mu\alpha\Gamma(\gamma_1)}{1 - \mu\alpha}$ got corrupt in $t = 0$, and since by Lemma 5 equilibrium sequences are strictly decreasing, we have

$$\begin{aligned}\gamma_1 &= \frac{1}{2} \left[\Gamma(\gamma_1) + F \left(\Gamma(\gamma_1); \frac{\Gamma(\gamma_0) - \mu\alpha\Gamma(\gamma_1)}{1 - \mu\alpha}, \mu \right) + E_F(\gamma_1, \gamma_2) \right] \\ \gamma'_1 &= \frac{1}{2} \left[\Gamma(\gamma'_1) + F \left(\Gamma(\gamma'_1); \frac{\Gamma(\gamma'_0) - \mu\alpha\Gamma(\gamma'_1)}{1 - \mu\alpha}, \mu \right) + E_F(\gamma'_1, \gamma'_2) \right]\end{aligned}$$

which, recalling (66), are equivalent to

$$\begin{aligned}2\gamma_1 - \frac{1}{1 - \mu\alpha}\Gamma(\gamma_1) + \frac{\mu\alpha}{1 - \mu\alpha}\Gamma(\gamma_2) &= F \left(\Gamma(\gamma_1); \frac{\Gamma(\gamma_0) - \mu\alpha\Gamma(\gamma_1)}{1 - \mu\alpha}, \mu \right) \\ 2\gamma'_1 - \frac{1}{1 - \mu\alpha}\Gamma(\gamma'_1) + \frac{\mu\alpha}{1 - \mu\alpha}\Gamma(\gamma'_2) &= F \left(\Gamma(\gamma'_1); \frac{\Gamma(\gamma'_0) - \mu\alpha\Gamma(\gamma'_1)}{1 - \mu\alpha}, \mu \right).\end{aligned}$$

Notice now that by (70) we have

$$\frac{\Gamma(\gamma_0) - \mu\alpha\Gamma(\gamma_1)}{1 - \mu\alpha} - \frac{\Gamma(\gamma'_0) - \mu\alpha\Gamma(\gamma'_1)}{1 - \mu\alpha} > 2(\gamma_0 - \gamma'_0)$$

and since by assumption $\gamma_0 > \gamma'_0$ we have that

$$\frac{\Gamma(\gamma_0) - \mu\alpha\Gamma(\gamma_1)}{1 - \mu\alpha} > \frac{\Gamma(\gamma'_0) - \mu\alpha\Gamma(\gamma'_1)}{1 - \mu\alpha}. \quad (73)$$

Given (72) and the fact that $\Gamma(\gamma)$ is decreasing, we also have that

$$\Gamma(\gamma'_1) > \Gamma(\gamma_1). \quad (74)$$

From (73) and (74) we then have

$$F \left(\Gamma(\gamma'_1); \frac{\Gamma(\gamma'_0) - \mu\alpha\Gamma(\gamma'_1)}{1 - \mu\alpha}, \mu \right) > F \left(\Gamma(\gamma_1); \frac{\Gamma(\gamma_0) - \mu\alpha\Gamma(\gamma_1)}{1 - \mu\alpha}, \mu \right)$$

which implies that

$$2\gamma'_1 - \frac{1}{1 - \mu\alpha}\Gamma(\gamma'_1) + \frac{\mu\alpha}{1 - \mu\alpha}\Gamma(\gamma'_2) > 2\gamma_1 - \frac{1}{1 - \mu\alpha}\Gamma(\gamma_1) + \frac{\mu\alpha}{1 - \mu\alpha}\Gamma(\gamma_2).$$

which is equivalent to

$$\Gamma(\gamma'_2) - \Gamma(\gamma_2) > 2\frac{1 - \mu\alpha}{\mu\alpha}(\gamma_1 - \gamma'_1) + \frac{1}{\mu\alpha}(\Gamma(\gamma'_1) - \Gamma(\gamma_1)) \quad (75)$$

and since both terms on the right hand side of inequality (75) are positive, and recalling that $\Gamma(\gamma)$ is decreasing, we have that

$$\gamma_2 > \gamma'_2. \quad (76)$$

Given (76) the argument above can be iterated to get

$$\Gamma(\gamma'_t) - \Gamma(\gamma_t) > 2\frac{1 - \mu\alpha}{\mu\alpha}(\gamma_{t-1} - \gamma'_{t-1}) + \frac{1}{\mu\alpha}(\Gamma(\gamma'_{t-1}) - \Gamma(\gamma_{t-1})) \quad (77)$$

for $t \geq 1$. Since

$$\begin{aligned} 2 \frac{1-\mu\alpha}{\mu\alpha} (\gamma_{t-1} - \gamma'_{t-1}) &> 0 \\ \Gamma(\gamma'_{t-1}) - \Gamma(\gamma_{t-1}) &> 0 \end{aligned}$$

and $\frac{1}{\mu\alpha} > 1$, by (77) $\Gamma(\gamma'_t) - \Gamma(\gamma_t)$ has to be increasing over time. By Lemma 5, however, equilibrium sequences converge to the steady state, so that $\lim_t (\Gamma(\gamma'_t) - \Gamma(\gamma_t)) = 0$, and a contradiction is obtained.

Similar arguments can be used to prove that given a distribution $G(\beta)$ with $G(\beta^*) \leq F(\beta^*, \beta^*, \mu)$ there cannot be two different equilibrium sequences starting from it. ■

A.5 Proof of Proposition 3

(1) Consider two sequences $\{\gamma_t\}_{t=0}^\infty$ and $\{\gamma'_t\}_{t=0}^\infty$ and suppose, contrary to the claim and without loss of generality, that

$$\gamma_0 > \gamma'_0 > \gamma^* \quad (78)$$

$$\gamma'_1 > \gamma_1 > \gamma^*. \quad (79)$$

Since we know from Lemma 5 that equilibrium sequences have to be strictly decreasing, equations (64) and (65) have to hold. It is easy to see that (64) and (65) are equivalent to

$$\Gamma(\gamma_2) = \frac{1-\mu\alpha}{\mu\alpha} F\left(\Gamma(\gamma_1); \frac{\Gamma(\gamma_0) - \mu\alpha\Gamma(\gamma_1)}{1-\mu\alpha}, \mu\right) - 2\frac{1-\mu\alpha}{\mu\alpha}\gamma_1 + \frac{1}{\mu\alpha}\Gamma(\gamma_1) \quad (80)$$

$$\Gamma(\gamma'_2) = \frac{1-\mu\alpha}{\mu\alpha} F\left(\Gamma(\gamma'_1); \frac{\Gamma(\gamma'_0) - \mu\alpha\Gamma(\gamma'_1)}{1-\mu\alpha}, \mu\right) - 2\frac{1-\mu\alpha}{\mu\alpha}\gamma'_1 + \frac{1}{\mu\alpha}\Gamma(\gamma'_1) \quad (81)$$

Subtracting (81) from (80) we get

$$\begin{aligned} \Gamma(\gamma_2) - \Gamma(\gamma'_2) &= \frac{1-\mu\alpha}{\mu\alpha} \left[F\left(\Gamma(\gamma_1); \frac{\Gamma(\gamma_0) - \mu\alpha\Gamma(\gamma_1)}{1-\mu\alpha}, \mu\right) - F\left(\Gamma(\gamma'_1); \frac{\Gamma(\gamma'_0) - \mu\alpha\Gamma(\gamma'_1)}{1-\mu\alpha}, \mu\right) \right] \\ &\quad - 2\frac{1-\mu\alpha}{\mu\alpha}(\gamma_1 - \gamma'_1) + \frac{1}{\mu\alpha}(\Gamma(\gamma_1) - \Gamma(\gamma'_1)) \end{aligned} \quad (82)$$

Since by (78) and (79)

$$\begin{aligned} \frac{1-\mu\alpha}{\mu\alpha} \left[F\left(\Gamma(\gamma_1); \frac{\Gamma(\gamma_0) - \mu\alpha\Gamma(\gamma_1)}{1-\mu\alpha}, \mu\right) - F\left(\Gamma(\gamma'_1); \frac{\Gamma(\gamma'_0) - \mu\alpha\Gamma(\gamma'_1)}{1-\mu\alpha}, \mu\right) \right] &> 0 \\ -2\frac{1-\mu\alpha}{\mu\alpha}(\gamma_1 - \gamma'_1) &> 0 \\ \frac{1}{\mu\alpha}(\Gamma(\gamma_1) - \Gamma(\gamma'_1)) &> 0 \end{aligned}$$

from (82) we have

$$\Gamma(\gamma_2) - \Gamma(\gamma'_2) > \frac{1}{\mu\alpha}(\Gamma(\gamma_1) - \Gamma(\gamma'_1)) > 0 \quad (83)$$

Consider now period $t = 2$. As above, from Lemma 5 we can show that the following two equations have to hold

$$\Gamma(\gamma_3) = \frac{1-\mu\alpha}{\mu\alpha} F\left(\Gamma(\gamma_2); \frac{\Gamma(\gamma_1) - \mu\alpha\Gamma(\gamma_2)}{1-\mu\alpha}, \mu\right) - 2\frac{1-\mu\alpha}{\mu\alpha}\gamma_2 + \frac{1}{\mu\alpha}\Gamma(\gamma_2) \quad (84)$$

$$\Gamma(\gamma'_3) = \frac{1-\mu\alpha}{\mu\alpha} F\left(\Gamma(\gamma'_2); \frac{\Gamma(\gamma'_1) - \mu\alpha\Gamma(\gamma'_2)}{1-\mu\alpha}, \mu\right) - 2\frac{1-\mu\alpha}{\mu\alpha}\gamma'_2 + \frac{1}{\mu\alpha}\Gamma(\gamma'_2) \quad (85)$$

Notice that, given (83) implies

$$\begin{aligned} \frac{\Gamma(\gamma'_1) - \mu\alpha\Gamma(\gamma'_2)}{1-\mu\alpha} &> \frac{\Gamma(\gamma_1) - \mu\alpha\Gamma(\gamma_2)}{1-\mu\alpha} \\ \gamma_2 - \gamma'_2 &< 0, \end{aligned}$$

we can argue as above that

$$\Gamma(\gamma_3) - \Gamma(\gamma'_3) > \frac{1}{\mu\alpha} (\Gamma(\gamma_2) - \Gamma(\gamma'_2)) > 0$$

and we can iterate the argument to get

$$\Gamma(\gamma_t) - \Gamma(\gamma'_t) > \frac{1}{\mu\alpha} (\Gamma(\gamma_{t-1}) - \Gamma(\gamma'_{t-1})) \quad (86)$$

for $t \geq 2$ which implies that $\Gamma(\gamma'_t) - \Gamma(\gamma_t)$ has to be increasing over time. By Lemma 5, however, equilibrium sequences converge to the steady state, so that $\lim_t (\Gamma(\gamma'_t) - \Gamma(\gamma_t)) = 0$, and a contradiction is obtained.

(2) Omitted.

(3) From Lemma 5 all equilibrium subsequences starting from $t = 1$ and different from infinite repetitions of the steady state level of corruption are above the steady state and are strictly decreasing.

Since $\gamma_t > \gamma^*$, $\beta_t^O = \Gamma(\gamma_t) < \Gamma(\gamma^*) = \beta^*$.

Since for $t = 1, 2, \dots$, $\gamma_t > \gamma_{t+1}$, we have $\Gamma(\gamma_t) < \Gamma(\gamma_{t+1})$ which implies

$$\beta_t^O = \Gamma(\gamma_t) < \Gamma(\gamma_{t+1}) = \beta_{t+1}^O.$$

For the same reason we also have

$$\beta_t^Y = \beta_t^{YC} = \frac{\Gamma(\gamma_t) - \mu\alpha\Gamma(\gamma_{t+1})}{1-\mu\alpha} < \Gamma(\gamma_t) = \beta_t^O.$$

which completes the proof. ■

A.6 Proof of Proposition 4

(1) From Lemma 5 we know that $\gamma_2 > \gamma_3 > \dots > \gamma^*$. We therefore only need to show that $\gamma_1 > \gamma^*$. To do this suppose that, contrary to the statement, $\gamma_1 < \gamma^*$. Given this implies that $\gamma_2 > \gamma_1$, the profitability

of corruption is higher at $t = 1$ than at $t = 2$, and this implies that all agents, both young and old, with types $\beta < \tilde{\Gamma}(\gamma_1)$ will get corrupt so that γ_1 will be determined by the following equation:

$$\gamma_1 = \frac{1}{2} \left[\tilde{\Gamma}(\gamma_1) + F(\tilde{\Gamma}(\gamma_1); \beta^*, \mu) \right]. \quad (87a)$$

If $\gamma_1 < \gamma^*$

$$\tilde{\Gamma}(\gamma_1) > \Gamma(\gamma_1) > \Gamma(\gamma^*) \quad (88)$$

and

$$F(\tilde{\Gamma}(\gamma_1); \beta^*, \mu) > F(\Gamma(\gamma^*); \beta^*, \mu). \quad (89)$$

It is now easy to see that (56), (88) and (89) provide a contradiction to (87a) when $\gamma_0 < \gamma^*$.

(2) From Lemma 5 we know that $\gamma_2 > \gamma_3 > \dots > \gamma^*$. We therefore only need to show that $\gamma_1 < \gamma^*$.

To do this suppose that, contrary to the statement, $\gamma_1 > \gamma^*$.

We have to consider two cases: (a) $\gamma_1 \leq \gamma_2$; (b) $\gamma_1 > \gamma_2$.

(a) A contradiction to the hypothesis that $\gamma_1 > \gamma^*$ is reached using the same arguments that were used in the proof of part 1, with all inequalities reversed.

(b) Since $\gamma_1 > \gamma_2 > \gamma^*$ the law of motion is

$$\gamma_1 = \frac{1}{2} \left[\tilde{\Gamma}(\gamma_1) + F(\tilde{\Gamma}(\gamma_1); \beta^*, \mu) + E_F(\gamma_1, \gamma_2) \right]. \quad (90)$$

Since $\gamma_1 > \gamma_2$, $E_F(\gamma_1, \gamma_2) < 0$ which, together with (90), implies

$$\gamma_1 < \frac{1}{2} \left[\tilde{\Gamma}(\gamma_1) + F(\tilde{\Gamma}(\gamma_1); \beta^*, \mu) \right]. \quad (91)$$

Under the assumption that $\gamma_1 > \gamma^*$ we have

$$\tilde{\Gamma}(\gamma_1) < \Gamma(\gamma_1) < \Gamma(\gamma^*) \quad (92)$$

and

$$F(\tilde{\Gamma}(\gamma_1); \beta^*, \mu) < F(\Gamma(\gamma^*); \beta^*, \mu). \quad (93)$$

and it is easy to see that (56), (92), and (93) provide a contradiction to (91) when $\gamma_1 > \gamma^*$. ■

A.7 Proof of Proposition 5

Since the change in μ is unforeseen until the end of period $t = 1$, we have that $\gamma_1 = \gamma^*$. Given this, and given μ' , the distribution of the agents' corruption costs will then be $F(\beta; \beta^*, \mu')$. Since it is easy to see that $F(\beta^*; \beta^*, \mu') > F(\beta^*; \beta^*, \mu)$ if $\mu' > \mu$ and $F(\beta^*; \beta^*, \mu') < F(\beta^*; \beta^*, \mu)$ if $\mu' < \mu$, applying Lemma 6 concludes the argument. ■

A.8 Proof of Proposition 6

To prove Proposition 6 we will first present a sequence of Lemmas.

Let $\beta^Y(\gamma)$ denote the mapping from equilibrium levels of corruption to the type of young agent who is indifferent between being honest and corrupt. Then we have:

Lemma 8 $\beta^Y(\gamma)$ is a strictly decreasing function.

Proof. The fact that $\beta^Y(\gamma)$ is a function is a straightforward consequence of Proposition 2 that shows that equilibrium sequences are unique. To show that $\beta^Y(\gamma)$ is decreasing, consider $\beta^Y(\gamma^*)$ and $\beta^Y(\gamma_t)$. By definition we have

$$\begin{aligned}\beta^Y(\gamma^*) &= \frac{\Gamma(\gamma^*) - \mu\alpha\Gamma(\gamma^*)}{1 - \mu\alpha} \\ \beta^Y(\gamma_t) &= \frac{\Gamma(\gamma_t) - \mu\alpha\Gamma(\gamma_{t+1})}{1 - \mu\alpha}\end{aligned}$$

Since $\gamma^* < \gamma_{t+1} < \gamma_t$, and since $\beta^Y(\gamma)$ is monotonic, to prove that $\beta^Y(\gamma)$ is decreasing it is sufficient to show that $\beta^Y(\gamma^*) > \beta^Y(\gamma_t)$. Suppose to the contrary that $\beta^Y(\gamma^*) < \beta^Y(\gamma_t)$. Simple algebra shows that this is equivalent to

$$\Gamma(\gamma^*) - \Gamma(\gamma_t) < \mu\alpha(\Gamma(\gamma^*) - \Gamma(\gamma_{t+1}))$$

which is impossible since Γ is decreasing and $\gamma^* < \gamma_{t+1} < \gamma_t$. ■

Given the results of Lemmas 6 and 8, the following Lemma shows that if it were known that two initial distributions of old agents generate equilibrium sequences whose first elements belong to an interval lying either above or below the steady state, a complete comparison of equilibrium sequences can be done, if one distribution dominates the other in an appropriate interval.

Lemma 9 Let two initial distributions of old agents at time $t = 0$ be described by the distribution functions $G'(\beta)$, and $G''(\beta)$ and let γ'_0 and γ''_0 denote the equilibrium corruption levels associated to them in $t = 0$. Suppose that both γ'_0 and γ''_0 belong to an interval $[\underline{\gamma}, \overline{\gamma}]$ with either $\gamma^* < \underline{\gamma} < \overline{\gamma}$ or $\underline{\gamma} < \overline{\gamma} < \gamma^*$. If $G''(\beta) > G'(\beta)$ for all $\beta \in [\Gamma(\overline{\gamma}), \Gamma(\underline{\gamma})]$, then $\gamma''_0 > \gamma'_0$.

Proof. First consider the case in which $\gamma^* < \underline{\gamma} < \overline{\gamma}$. In this case, by Lemma 6, γ'_0 is given by the following equation

$$\gamma'_0 = \frac{1}{2} \left[G'(\Gamma(\gamma'_0)) + \Gamma(\gamma'_0) - \frac{\mu\alpha}{1 - \mu\alpha} [\Gamma(\gamma'_1) - \Gamma(\gamma'_0)] \right]. \quad (94)$$

Let

$$h(\gamma'_0) = 2\gamma'_0 - \Gamma(\gamma'_0) + \frac{\mu\alpha}{1 - \mu\alpha} [\Gamma(\gamma'_1) - \Gamma(\gamma'_0)] = 2\gamma'_0 - \beta^Y(\gamma'_0) \quad (95)$$

and notice that, since by Lemma 8 $\beta^Y(\cdot)$ is decreasing, $h(\cdot)$ is increasing. Using (95) we can rewrite (94) as

$$h(\gamma'_0) = G'(\Gamma(\gamma'_0)). \quad (96)$$

Similarly γ''_0 is given by the following equation

$$h(\gamma''_0) = G''(\Gamma(\gamma''_0)). \quad (97)$$

Subtract (96) from (97) to get

$$h(\gamma''_0) - h(\gamma'_0) = G''(\Gamma(\gamma''_0)) - G'(\Gamma(\gamma'_0)). \quad (98)$$

Suppose, contrary to the statement, that $\gamma''_0 < \gamma'_0$. This implies that the left hand side of (98) is negative and the right hand side positive and therefore leads to a contradiction.

Now consider the case in which $\underline{\gamma} < \overline{\gamma} < \gamma^*$. In this case γ'_0 is given by the following equation

$$\gamma'_0 = \frac{1}{2} [\Gamma(\gamma'_0) + G'(\Gamma(\gamma'_0))]. \quad (99)$$

Let

$$\tilde{h}(\gamma'_0) = 2\gamma'_0 - \Gamma(\gamma'_0) \quad (100)$$

and notice that, since $\Gamma(\cdot)$ is decreasing, $\tilde{h}(\cdot)$ is increasing. Using (100) we can rewrite (99) as

$$\tilde{h}(\gamma'_0) = G'(\Gamma(\gamma'_0)).$$

Similarly γ''_0 is given by the following equation

$$\tilde{h}(\gamma''_0) = G''(\Gamma(\gamma''_0)).$$

The argument following (97) completes the proof. ■

The following Lemma establishes a result that will be extensively used in the following to ascertain which of conditions (54) and (55) of Lemmas 6 holds when studying comparative dynamics.

Lemma 10 *Let β' and β'' be given.*

1. *If $\beta' < \beta''$, then $F(\beta''; \beta', \mu') > F(\beta''; \beta'', \mu'')$ if and only if $\mu' \beta' < \mu'' \beta''$.*
2. *If $\beta' > \beta''$, then $F(\beta''; \beta', \mu') > F(\beta''; \beta'', \mu'')$ if and only if $\mu' (1 - \beta') < \mu'' (1 - \beta'')$.*

Proof. When $\beta' < \beta''$

$$\begin{aligned} F(\beta''; \beta', \mu') &= (1 - \mu' + \mu' \beta') \beta' + (1 + \mu' \beta') (\beta'' - \beta') > \\ &> (1 - \mu'' + \mu'' \beta'') \beta'' = F(\beta''; \beta'', \mu''). \end{aligned}$$

Rearranging gives

$$-\mu' \beta' (1 - \beta'') > -\mu'' \beta'' (1 - \beta'')$$

or

$$\mu' \beta' < \mu'' \beta''.$$

When $\beta' > \beta''$

$$\begin{aligned} F(\beta''; \beta', \mu') &= (1 - \mu' + \mu' \beta') \beta'' > \\ &> (1 - \mu'' + \mu'' \beta'') \beta'' = F(\beta''; \beta'', \mu'') \end{aligned}$$

and rearranging gives

$$-\mu' \beta'' (1 - \beta') > -\mu'' \beta'' (1 - \beta'')$$

or

$$\mu' (1 - \beta') < \mu'' (1 - \beta'').$$

■

The following Lemma studies the equilibrium sequence starting from a given initial distribution of old agents $F(\beta; \beta', \mu')$, with $\mu' = \mu$. Using the results of Lemmas 6, 9, and 10 it shows that equilibrium sequences always start above the steady state and that the starting point is decreasing in β' when $\beta' < \beta^*$ and increasing in β' when $\beta' > \beta^*$.

Lemma 11 *Given an initial distribution of old agents $F(\beta; \beta', \mu)$ at time $t = 0$ the equilibrium sequence $\{\gamma'_t\}_{t=0}^\infty$ is such that*

1. $\hat{\gamma} > \gamma'_0 > \gamma'_1 > \gamma'_2 > \dots > \gamma^*$.
2. When $\beta' < \beta^*$, γ'_0 is a decreasing function of β' .
3. When $\beta' > \beta^*$, γ'_0 is an increasing function of β' .

Proof. (1) Consider $F(\beta; \beta', \mu)$ and $F(\beta; \beta^*, \mu)$. Then, using Lemma 10 it is easy to see that for both $\beta' < \beta$ and $\beta' > \beta$

$$F(\beta^*; \beta', \mu) > F(\beta^*; \beta^*, \mu).$$

Since this is equivalent to (54), the result that $\gamma'_0 > \gamma'_1 > \gamma'_2 > \dots > \gamma^*$ follows from Lemma 6. The proof of the fact that $\hat{\gamma} > \gamma'_0$ will be delayed until after the proof of parts (2) and (3).

(2) Consider two distributions of old agents $F(\beta; \beta', \mu)$ and $F(\beta; \beta'', \mu)$ with $\beta' < \beta''$ and let $\{\gamma'_t\}_{t=0}^\infty$ and $\{\gamma''_t\}_{t=0}^\infty$ denote the associated equilibrium sequences of corruption starting from such initial distributions. From above we know that $\gamma'_0 > \gamma^*$ and $\gamma''_0 > \gamma^*$ which implies that $\Gamma(\gamma'_0) < \beta^*$ and $\Gamma(\gamma''_0) < \beta^*$.

Now suppose that $\beta' < \beta'' < \beta^*$. We want to establish lower bounds on $\Gamma(\gamma'_0)$ and $\Gamma(\gamma''_0)$ that combined with the upper bound provided above will enable us to apply Lemma 9. To do this let $\tilde{\beta}(\beta'')$ be the value of β implicitly defined by the following equation (see figure 6)

$$F(\tilde{\beta}(\beta''); \beta'', \mu) = F(\beta^*; \beta^*, \mu).$$

Notice that since

$$\begin{aligned} F(\beta^*; \beta'', \mu) &= (1 - \mu + \mu\beta'')\beta'' + (1 + \mu\beta'')(\beta^* - \beta'') \\ &> (1 - \mu + \mu\beta^*)\beta^* = F(\beta^*; \beta^*, \mu) \end{aligned}$$

$\tilde{\beta}(\beta'') < \beta^*$ and since $F(\beta; \beta'', \mu) < F(\beta; \beta^*, \mu)$, for $\beta \in [0, \beta'']$, $\tilde{\beta}(\beta'') > \beta''$ (see figure 6 again). Now suppose that $\Gamma(\gamma''_0) < \tilde{\beta}(\beta'')$. This implies that the mass of corrupt old agents is less than in the steady state ($F(\beta^*; \beta^*, \mu)$) which implies that to achieve a $\gamma''_0 > \gamma^*$ the mass of corrupt young agents in 0 has to be higher than in the steady state. Since this is impossible we conclude $\Gamma(\gamma''_0) > \tilde{\beta}(\beta'')$. Now for given β', β'' , with $\beta' < \beta'' < \beta^*$, let $\hat{\beta}$ be such that $F(\hat{\beta}; \beta', \mu) = F(\hat{\beta}; \beta'', \mu)$ (see figure 6) and suppose that $\Gamma(\gamma'_0) < \hat{\beta}$. Since, as figure 6 shows, $\hat{\beta} \in (\beta', \beta'')$ and $F(\beta; \beta', \mu) < F(\beta; \beta'', \mu)$ for all $\beta < \hat{\beta}$, $\Gamma(\gamma'_0) < \hat{\beta}$ implies that the mass of old agents corrupt in 0 is less than in the steady state, $F(\Gamma(\gamma'_0); \beta', \mu) < F(\Gamma(\gamma'_0); \beta'', \mu) < F(\beta^*; \beta^*, \mu)$, which implies that, in order to have $\gamma'_0 > \gamma^*$, the mass of corrupt young agents in 0 has to be larger than in the steady state. Since this is impossible we conclude that $\Gamma(\gamma'_0) > \hat{\beta}$.

We now know that $\Gamma(\gamma'_0) \in (\hat{\beta}, \beta^*)$ and $\Gamma(\gamma''_0) \in (\tilde{\beta}(\beta''), \beta^*)$ which implies that $\Gamma(\gamma''_0) \in (\hat{\beta}, \beta^*)$ (since $\tilde{\beta}(\beta'') > \beta'' > \hat{\beta}$). Recalling that $F(\beta; \beta', \mu) > F(\beta; \beta'', \mu)$ for all $\beta \in (\hat{\beta}, \beta^*)$ we can apply the result of Lemma 9 which implies that $\gamma'_0 > \gamma''_0$ and concludes the proof of Part 2.

(3) Consider two distributions of old agents $F(\beta; \beta', \mu)$ and $F(\beta; \beta'', \mu)$ with $\beta' > \beta''$ and let $\{\gamma'_t\}_{t=0}^\infty$ and $\{\gamma''_t\}_{t=0}^\infty$ denote the associated equilibrium sequences of corruption starting from such initial distributions. From above we know that $\gamma'_0 > \gamma^*$ and $\gamma''_0 > \gamma^*$ which implies that $\Gamma(\gamma'_0) < \beta^*$ and $\Gamma(\gamma''_0) < \beta^*$. Now suppose that $\beta' > \beta'' > \beta^*$; then from Lemma 9 we know that a sufficient condition for $\gamma'_0 > \gamma''_0$ is that $F(\beta; \beta', \mu) > F(\beta; \beta'', \mu)$ for all $\beta \in (0, \beta^*)$. Since for all $\beta \in [0, \beta^*]$

$$F(\beta; \beta', \mu) = (1 - \mu + \mu\beta')\beta > (1 - \mu + \mu\beta'')\beta = F(\beta; \beta'', \mu)$$

we have

$$F(\beta; \beta', \mu) > F(\beta; \beta'', \mu)$$

for all $\beta \in [0, \beta^*]$ which concludes the proof of part 3.

Using the previous results we now want to show that given an initial distribution of old agents $F(\beta; \beta', \mu)$ at time 0 the equilibrium sequence $\{\gamma'_t\}_{t=0}^\infty$ is such that $\gamma'_0 < \hat{\gamma}$ (the level of corruption in the static model). When $\beta' < \beta^*$ we know from part 2 that γ'_0 is decreasing in β' . To show that $\gamma'_0 < \hat{\gamma}$ for all $\beta' \in [0, \beta^*]$ it is therefore sufficient to show that $\gamma'_0 < \hat{\gamma}$ when $\beta' = 0$. Since we have already shown that $\gamma'_0 > \gamma'_1 > \gamma^*$ the level of corruption at time 0 will be

$$\gamma'_0 = \frac{1}{2} \left[\beta^{YC}(\gamma'_0, \gamma'_1) + F(\Gamma(\gamma'_0); \bar{\beta}, \mu) \right]$$

which can be shown to be equivalent to

$$\gamma'_0 = \Gamma(\gamma'_0) + \frac{\mu\alpha(\Gamma(\gamma'_0) - \Gamma(\gamma'_1))}{2(1 - \mu\alpha)} \quad (101)$$

Recall from equation (8) that $\hat{\gamma}$ is such that

$$\hat{\gamma} = \Gamma(\hat{\gamma}). \quad (102)$$

Subtracting (102) from (8) we get

$$\gamma'_0 - \hat{\gamma} = \Gamma(\gamma'_0) - \Gamma(\hat{\gamma}) + \frac{\mu\alpha(\Gamma(\gamma'_0) - \Gamma(\gamma'_1))}{2(1 - \mu\alpha)}. \quad (103)$$

Suppose that $\gamma'_0 > \hat{\gamma}$. Then recalling that $\gamma'_0 > \gamma'_1$ and that Γ is decreasing both terms on the right hand side of (103) are positive and we get a contradiction.

When $\beta' > \beta^*$ we know from part 2 that γ'_0 is increasing in β' . To show that $\gamma'_0 < \hat{\gamma}$ (the level of corruption in the static model) for all $\beta' \in [\beta^*, 1]$ it is therefore sufficient to show that $\gamma'_0 < \hat{\gamma}$ when $\beta' = 1$. Since we have already shown that $\gamma'_0 > \gamma'_1 > \gamma^*$ the level of corruption will be given by equation (101) and the same argument as above applies. ■

Given the above results we can now move to the proof of Proposition 6.

(1) Since $\gamma' < \gamma^*$, it is easy to see that given that $\beta' < \beta^*$, the distribution function of old agents at date 0 is $F(\beta; \bar{\beta}, \mu)$ with $\bar{\beta} = \beta' < \beta^*$. Part 1a follows directly applying part 1 of Lemma 11. Part 1b follows applying part 2 of Lemma 11. (2) Since $\gamma' > \gamma^*$, it is easy to see that given that $\beta' > \beta^*$, the distribution function of old agents at date 0 is $F(\beta; \bar{\beta}, \mu)$ with $\bar{\beta} = \beta' > \beta^*$. Part 2a follows directly applying part 1 of Lemma 11. Part 2b follows applying part 3 of Lemma 11. ■

A.9 Permanent changes in the detection probability

When considering permanent changes in μ , it is important first to determine whether changes in μ imply increases or decreases in β^* , the type of agent who is just indifferent between getting corrupt or not in the steady state. In the rest of the paper we will make the following Assumption.

Assumption 1 $\partial\beta^*/\partial\mu < 0$.

This assumption is justified once we recognize that in the case in which $V(q) = \log q$ and $C(q) = \theta q$ one can show the following result

Lemma 12 *Suppose that $V(q) = \log q$ and $C(q) = \theta q$. Then $B_A + B_F > 1/8$ is a sufficient condition for $\partial\beta^*/\partial\mu < 0$.*

Proof. Recall that

$$\beta^* = 1 - \gamma^*(\mu) - k = 1 - \gamma^*(\mu) - \theta q_c - \mu(B_A + B_F).$$

Differentiating with respect to μ , we get

$$\frac{\partial\beta^*}{\partial\mu} = -\frac{\partial\gamma^*(\mu)}{\partial\mu} - \frac{\partial k}{\partial\mu} = -\frac{\partial\gamma^*(\mu)}{\partial\mu} - (B_A + B_F). \quad (104)$$

Recall that γ^* is given by the smaller root of

$$H = \mu(\gamma^*)^2 - [4 + \mu - 2\mu k]\gamma^* + (2 - \mu)(1 - k) + \mu(1 - k)^2 = 0.$$

Applying the implicit function theorem we get

$$\frac{\partial\gamma^*(\mu)}{\partial\mu} = -\frac{\frac{\partial H}{\partial\mu} + \frac{\partial H}{\partial k} \frac{\partial k}{\partial\mu}}{\frac{\partial H}{\partial\gamma^*}}.$$

Straightforward, although tedious calculations show that substituting the result in (104) we get

$$\frac{\partial\beta^*}{\partial\mu} = \frac{(\gamma^* + k)(1 - \gamma^* - k) - 2(B_A + B_F)}{4 + \mu - 2\mu k - 2\mu\gamma^*}.$$

Given the denominator is positive, $\frac{\partial\beta^*}{\partial\mu} < 0$ if and only if

$$B_A + B_F > \frac{(\gamma^* + k)(1 - \gamma^* - k)}{2}. \quad (105)$$

Since

$$\beta^* = 1 - \gamma^* - k$$

(105) can be rewritten as

$$B_A + B_F > \frac{(1 - \beta^*)\beta^*}{2}. \quad (106)$$

Since $\beta^* \in [0, 1]$

$$\frac{(1 - \beta^*)\beta^*}{2} < \frac{1}{8}$$

and a sufficient condition for (106) to hold is

$$B_A + B_F > \frac{1}{8}. \quad (107)$$

■

Lemma 12 shows that if the sum of the penalties imposed on the agent and the contractor is not too small, then $\partial\beta^*/\partial\mu < 0$. It is probably worth noticing that the condition of Lemma 12 does not seem very restrictive if we insist on the fact that penalties should at least be such that no agent would want to be corrupt if he knew that he was going to be caught with probability 1. A sufficient condition in order to rule out such possibility is that $0 > \theta q - \theta q_c - (B_A + B_F)$ (an upper bound on the aggregate payoff to the agent and the contractor when $\mu = 1$) which is equivalent to $B_A + B_F > \theta q - \theta q_c$. Given that for $B_A + B_F > 1/8$ to hold, it is sufficient that $q_c < q - \frac{1}{8\theta} < \frac{7}{8\theta} = \frac{7}{8} \sup q$ where $\sup q = 1/\theta$ is an upper bound on the equilibrium level of desired quality as it is the principal's desired level of quality with zero corruption.

Given Assumption 1, we are now in the position to state the following Proposition characterizing comparative dynamics following a change in μ .

Proposition 7 *Let γ' and γ'' be the steady states associated to μ' and μ'' respectively. Suppose that Assumption 1 is satisfied, let γ^* denote the steady state associated to μ^* , and let $\{\gamma'_t\}_{t=1}^\infty$ and $\{\gamma''_t\}_{t=1}^\infty$ denote the equilibrium sequences starting at $t = 1$ given that corruption at time $t = 0$ was in steady state γ' and γ'' , respectively. Then, a permanent change to μ^* has the following implications:*

1. *If $\gamma' < \gamma^*$ ($\Leftrightarrow \mu' > \mu^*$) and $\mu^*\beta^* > \mu'\beta'$ the equilibrium sequence $\{\gamma'_t\}_{t=1}^\infty$ will be such that $\hat{\gamma} > \gamma'_1 > \gamma'_2 > \gamma'_3 > \dots > \gamma^*$.*
2. *If $\gamma' < \gamma'' < \gamma^*$ ($\Leftrightarrow \mu' > \mu'' > \mu^*$) and $\mu^*\beta^* > \mu''\beta'' > \mu'\beta'$ the equilibrium sequences $\{\gamma'_t\}_{t=1}^\infty$ and $\{\gamma''_t\}_{t=1}^\infty$ will be such that $\gamma'_t > \gamma''_t$ for $t = 1, 2, \dots$.*
3. *If $\gamma' < \gamma^*$ ($\Leftrightarrow \mu' > \mu^*$) and $\mu^*\beta^* < \mu'\beta'$ the equilibrium sequence $\{\gamma'_t\}_{t=1}^\infty$ will be such that $\gamma'_1 < \gamma^*$ and $\gamma'_2 > \gamma'_3 > \dots > \gamma^*$.*
4. *If $\gamma' < \gamma'' < \gamma^*$ ($\Leftrightarrow \mu' > \mu'' > \mu^*$) and $\mu^*\beta^* < \mu''\beta'' < \mu'\beta'$ the equilibrium sequences $\{\gamma'_t\}_{t=1}^\infty$ and $\{\gamma''_t\}_{t=1}^\infty$ will be such that $\gamma'_1 < \gamma''_1$ and $\gamma'_t > \gamma''_t$ for $t = 2, 3, \dots$.*
5. *If $\gamma' > \gamma^*$ ($\Leftrightarrow \mu' < \mu^*$) the equilibrium sequence $\{\gamma'_t\}_{t=1}^\infty$ will be such that $\hat{\gamma} > \gamma'_1 > \gamma'_2 > \gamma'_3 > \dots > \gamma^*$.*
6. *If $\gamma'' > \gamma' > \gamma^*$ ($\Leftrightarrow \mu^* > \mu' > \mu''$) the equilibrium sequences $\{\gamma'_t\}_{t=1}^\infty$ and $\{\gamma''_t\}_{t=1}^\infty$ will be such that $\gamma'_t < \gamma''_t$ for all $t = 1, 2, \dots$.*

Proof. (1) Consider the distribution of old agents in the steady state γ' , $F(\beta'; \beta', \mu')$. Since, by Assumption 1, $\beta' < \beta^*$, from Lemma 10

$$F(\beta^*; \beta', \mu') > F(\beta^*; \beta^*, \mu^*) \quad (108)$$

if and only if $\mu'\beta' < \mu^*\beta^*$. Since (108) is equivalent to (54), $\gamma'_1 > \gamma'_2 > \dots > \gamma^*$ follows from Lemma 6. The proof of the fact that $\hat{\gamma} > \gamma'_1$ is a straightforward repetition of the argument used in the Proof of Lemma 11 and will therefore be omitted both here and in part 5 below.

(2) Consider the old agent steady state distributions associated to μ' and μ'' , with $\mu' > \mu'' > \mu^*$, $F(\beta; \beta', \mu')$, $F(\beta; \beta'', \mu'')$. Under Assumption 1 we know that $\beta' < \beta'' < \beta^*$ and, from above, $\gamma'_1 > \gamma^*$ and $\gamma''_1 > \gamma^*$ which implies that $\Gamma(\gamma'_1) < \beta^*$ and $\Gamma(\gamma''_1) < \beta^*$. We want to establish lower bounds on $\Gamma(\gamma'_1)$ and $\Gamma(\gamma''_1)$ that combined with the upper bound provided above will enable us to apply Lemma 9. To do this let $\tilde{\beta}(\beta'')$ be the value of β implicitly defined by the following equation

$$F(\tilde{\beta}(\beta''); \beta'', \mu'') = F(\beta^*; \beta^*, \mu^*)$$

Since Assumption 1 and Lemma 10 imply

$$F(\beta^*; \beta'', \mu'') > F(\beta^*; \beta^*, \mu^*),$$

we know that $\tilde{\beta}(\beta'') \in (\beta'', \beta^*)$ (see figure 7). Suppose that $\Gamma(\gamma''_1) < \tilde{\beta}(\beta'')$. This implies that the mass of corrupt old agents is less than in the steady state ($F(\beta^*; \beta^*, \mu^*)$) which implies that to achieve a $\gamma''_1 > \gamma^*$ the mass of corrupt young agents in 1 has to be higher than in the steady state. Since this is impossible we conclude $\Gamma(\gamma''_1) > \tilde{\beta}(\beta'')$. Now for given β' , β'' , with $\beta' < \beta'' < \beta^*$, let $\hat{\beta}$ be such that $F(\hat{\beta}; \beta', \mu') = F(\hat{\beta}; \beta'', \mu'')$ and suppose that $\Gamma(\gamma'_1) < \hat{\beta}$. Since Assumption 1 and Lemma 10 imply that $\hat{\beta} \in (\beta', \beta'')$ and since $F(\beta; \beta', \mu') < F(\beta; \beta'', \mu'')$ for all $\beta < \hat{\beta}$, $\Gamma(\gamma'_1) < \hat{\beta}$ implies that the mass of old agents corrupt in 1 is less than in the steady state, $F(\Gamma(\gamma'_1); \beta', \mu') < F(\Gamma(\gamma'_1); \beta'', \mu'') < F(\beta^*; \beta^*, \mu)$, which implies that, in order to have $\gamma'_1 > \gamma^*$, the mass of corrupt young agents in 1 has to be larger than in the steady state. Since this is impossible we conclude that $\Gamma(\gamma'_1) > \hat{\beta}$.

We now know that $\Gamma(\gamma'_1) \in (\hat{\beta}, \beta^*)$ and $\Gamma(\gamma''_1) \in (\tilde{\beta}(\beta''), \beta^*)$ which implies that $\Gamma(\gamma''_1) \in (\hat{\beta}, \beta^*)$ (since $\tilde{\beta}(\beta'') > \beta'' > \hat{\beta}$). Recalling that $F(\beta; \beta', \mu) > F(\beta; \beta'', \mu)$ for all $\beta \in (\hat{\beta}, \beta^*)$ we can apply the result of Lemma 9 which implies that $\gamma'_1 > \gamma''_1$ thus proving 2.

(3) Since $\mu^*\beta^* < \mu'\beta'$, from Lemma 10 we know that

$$F(\beta^*; \beta', \mu') < F(\beta^*; \beta^*, \mu^*)$$

which is equivalent to (55) which implies that part 3 follows directly from Lemma 6.

(4) consider the old agent steady state distributions associated to μ' and μ'' , with $\mu' > \mu'' > \mu^*$, $F(\beta; \beta', \mu')$, $F(\beta; \beta'', \mu'')$. It is easy to see that $F(\beta; \beta'', \mu'') > F(\beta; \beta', \mu')$, for all $\beta \in [0, 1]$. To see this first notice that, under Assumption 1, $\beta' < \beta'' < \beta^*$ and that the density associated to $F(\beta; \beta', \mu')$ in $[0, \beta']$ is strictly less than the density associated to $F(\beta; \beta'', \mu'')$ in the same interval, therefore implying that $F(\beta; \beta'', \mu'') > F(\beta; \beta', \mu')$, for all $\beta \in [0, \beta']$. Given both $F(\beta; \beta', \mu')$ and $F(\beta; \beta'', \mu'')$ are linear on $[\beta', \beta'']$, they can cross at most once in this interval. Given $\mu''\beta'' < \mu'\beta'$, however, from Lemma 10 $F(\beta''; \beta'', \mu'') > F(\beta''; \beta', \mu')$, which implies that $F(\beta; \beta'', \mu'') > F(\beta; \beta', \mu')$ for $\beta \in [\beta', \beta'']$. Finally, given both $F(\beta; \beta', \mu')$ and $F(\beta; \beta'', \mu'')$ are linear on $[\beta'', 1]$, they can cross at most once in this interval, but since $F(1; \beta'', \mu'') = F(1; \beta', \mu')$ we get $F(\beta; \beta'', \mu'') > F(\beta; \beta', \mu')$ for $\beta \in [\beta'', 1]$ (see figure 8). Since $F(\beta; \beta'', \mu'') > F(\beta; \beta', \mu')$, for all $\beta \in [0, 1]$, we can apply Lemma 9 which implies that $\gamma'_1 < \gamma''_1$. Given this and noticing that the old agents' equilibrium distributions in 2 will be respectively $F(\beta; \Gamma(\gamma'_1), \mu^*)$ and $F(\beta; \Gamma(\gamma''_1), \mu^*)$ and that both $\Gamma(\gamma'_1) > \beta^*$, and $\Gamma(\gamma''_1) > \beta^*$ the fact that γ_t is decreasing in γ' for all $t > 1$ follows directly from part 2 of Lemma 11.

(5) Since, by Assumption 1 $\beta' > \beta^*$, from Lemma 10

$$F(\beta^*; \beta', \mu') > F(\beta^*; \beta^*, \mu^*) \quad (109)$$

if and only if $\mu'(1 - \beta') < \mu^*(1 - \beta^*)$. Since $\beta' > \beta^*$ implies $1 - \beta' < 1 - \beta^*$, the latter, together with $\mu' < \mu^*$, implies that $\mu'(1 - \beta') < \mu^*(1 - \beta^*)$ and therefore that (109) holds. Since (109) is equivalent to (54), part 5 follows directly from Lemma 6.

(6) Consider the old agent steady state distributions associated to μ' and μ'' , with $\mu^* > \mu' > \mu''$, $F(\beta; \beta', \mu')$, $F(\beta; \beta'', \mu'')$. Given the above ($\gamma'_1 > \gamma^*$, $\gamma''_1 > \gamma^*$) we have $\Gamma(\gamma'_1) < \beta^*$, $\Gamma(\gamma''_1) < \beta^*$. Since $\beta'' > \beta' > \beta^*$, it is easy to see that since the densities associated to $F(\beta; \beta', \mu')$ and $F(\beta; \beta'', \mu'')$ on $[0, \beta^*]$ are respectively $1 - \mu'(1 - \beta')$ and $1 - \mu''(1 - \beta'')$ and that since $\mu'(1 - \beta') > \mu''(1 - \beta'')$ we have $F(\beta; \beta', \mu') < F(\beta; \beta'', \mu'')$ for all $\beta \in [0, \beta^*]$. Applying Lemma 9 concludes the Proof. ■

Proposition 7 generalizes the result of Proposition 6 to cases in which the change in the steady state is determined by a change in the probability with which a corrupt agent is detected and therefore fired. The results discussed in items 1 and 2 on page 23 also apply to this case with only one minor change in item 2a: When $\mu' > \mu^*$ (which implies that the old steady state is lower than the new one, $\gamma' < \gamma^*$) and $\mu^*\beta^* < \mu'\beta'$ the first level of corruption after the change from μ' to μ^* is lower than new steady state and is moreover *higher* the *higher* the old steady state; Nevertheless the monotonicity properties in γ' of the rest of the elements of the equilibrium sequence are unchanged, in the sense that the level of corruption at all other dates after the first will be *higher* the *lower* the old steady state.

References

- [1] Andvig, J. C. and K. O. Moene (1990) “How corruption may corrupt”, *Journal of Economic Behavior and Organization*, 13, 63-76.
- [2] Banfield (1975) “Corruption as a feature of governmental organization”, *Journal of Law and Economics*, 58, 587-605.
- [3] Bardhan, P. (1997) “Corruption and development: A review of issues”, *Journal of Economic Literature*, 35(3), 1320-1346.
- [4] Celarier, M. (1997) “Privatization: A case study in corruption”, *Journal of International Affairs*, 50(2), 531-43.
- [5] Celentani, M. and J.-J. Ganuza (1997) “Auctions and the dynamics of corruption”, GREMAQ, Cahier n. 97.16.459.
- [6] Lui, F. T. (1986) “A dynamic model of corruption”, *Journal of Public Economics*, 31, 215-36.
- [7] Rose-Ackerman, S. (1975) “The economics of corruption”, *Journal of Public Economics*, 4, 187-203.
- [8] Rose-Ackerman, S. (1978), *Corruption, a study in political economy*, Academic Press, New York, NY.
- [9] Shleifer, A. and R. W. Vishny (1993) “Corruption”, *Quarterly Journal of Economics*, 108, 599-617.
- [10] Stigler, G. (1970), “The optimum enforcement of law”, *Journal of Political Economy*, 78, 526-536.
- [11] *The Economist* (1997) “A (slightly) new Italian establishment”, 12 July 1997, 48.
- [12] Tirole, J. (1993) “A theory of collective reputations (with applications to the persistence of corruption and to firm quality)”, *Review of Economic Studies*, 63(1), 1-22.
- [13] Twain, M. (1899) *The man that corrupted Hadleyburg*, reprint in Penguin Books, London, 1995.

Figure 1: Equilibrium in the static setting

Figure 2: Old agents' density function

Figure 3: Old agents' distribution function

Figure 4: Old agents' distribution functions (for given μ and different β 's)

FORMER MOTTO

REVISED MOTTO

Figure 5: From Mark Twain (1899), page 59.

Figure 6: Old agents' distribution functions (for given μ and different β 's)

Figure 7: Old agents' distribution functions (for different μ 's and β 's)

Figure 8: Old agents' distribution functions (for different μ 's and β 's)