ON THE ECONOMICS OF WHISTLE-BLOWING BEHAVIOUR: THE ROLE OF INCENTIVES

Villena, Mauricio G. and Villena, Marcelo J.

School of Business, Adolfo Ibáñez University, Chile

24 March 2010

ON THE ECONOMICS OF WHISTLE-BLOWING BEHAVIOUR: THE ROLE OF INCENTIVES

by

Mauricio G. Villena† and Marcelo J. Villena‡

† *Corresponding Author*. School of Business, Adolfo Ibáñez University, Chile. Av. Diagonal las Torres 2640, Peñalolén. Santiago, Chile. Phone: 56/2/3311492, Fax: 56/2/2784413, e-mail: mauricio.villena@uai.cl.
‡ School of Engineering, Adolfo Ibáñez University, Chile

Abstract

The role of whistle-blowing as a mechanism for deterring corruption has been conspicuously neglected in the economic literature. This is quite surprising given the increase in legislation aimed at preventing corruption that includes whistle-blowing clauses and the extensive literature on whistle-blowing outside economics. In fact, we know of no formal economic model that deals squarely with the analysis of the role and potential impact of whistle-blowing on the persistence of corruption in organizations. Therefore, in an attempt to at least partially fill this gap, we present a theoretical model for approaching the issue, focusing specifically on the role of economic incentives to encourage whistle-blowing behaviour. We model corruption as a social norm of behaviour using elements of evolutionary game theory (EGT). We use the concept of replicator dynamics to explore the local asymptotical stability of several types of behaviour within organizations: (i) honest, corrupt, and honest whistle-blowing and (ii) honest, corrupt whistle-blowing, and honest whistle-blowing.

Keywords: Corruption, whistle-blowing, social norms, evolutionary games
JEL classification: K42, D73, C73

On The Economics of Whistle-Blowing Behaviour: The Role of Incentives

1. Introduction

Recently, the economics of corruption has given rise to a vast body of theoretical and empirical literature.[1] However, the role of *whistle-blowing* as a deterrence mechanism against corruption has been conspicuously neglected in the literature. This is quite surprising regarding the increase in legislation aimed at preventing corruption that includes whistle-blowing clauses[2] and the extensive literature on whistle-blowing outside economics.[3] We know of no formal economic model that deals squarely with the analysis of the role and potential impact of whistle-blowing on the *persistence of corruption* in organizations. Therefore, in an attempt to at least partially fill this gap, we present a theoretical model for approaching the issue, focusing specifically on the role of economic incentives aimed at encouraging whistle-blowing behaviour.

A quick look at a real life case study will help clarify the nature of whistle-blowing and our modelling strategy. In Canada in the mid-1990s, the Liberal Party implemented a federal "sponsorship programme" in Quebec Province in an attempt to increase awareness of the government's contributions to Quebec and discourage its separatism, as promoted by the provincial government (Parti Québécois). This programme lasted from 1996 until 2004, when corruption was exposed and its operations came under investigation by the Gomery Commission[4]. The allegations of corruption in the programme management referred mainly to payments of commissions for no apparent services and improper advances made to agencies related to the Liberal Party. These accusations focused on firms favoured by the sponsorship programme that maintained Liberal organizers or fundraisers on their payrolls or donated back part of the money to the Liberal Party. For years, this was an ongoing affair until Allan Cutler, a civil servant working for the Ministry of Public Works and Government Services, lodged a complaint that prompted a departmental audit of the advertising and public opinion division. Cutler was the *whistle-blower* who detected and reported the anomalies in the Canadian sponsorship programme and his actions triggered the so called "Sponsorgate" or "AdScam" scandal.

---

[1] For a survey of theoretical works see: Aidt (2003); for empirical works see: Bardhan (1997) and Jain (2001).
[2] See, for example, Groeneweg (2001) for a review of the distinct whistle-blower protection models of Australia, the United Kingdom, and the United States.
[3] See, for instance, Elliston (1985), Glazer and Glazer (1989), Jos and Tompkins (1989), Miceli and Near (1992), Clark (1997), Hunt (1997), Jubb (1999), Miethe (1999), and Alford (2001).
[4] Formally: the Commission of Inquiry into the Sponsorship Programme and Advertising Activities.

Some interesting questions arise from this specific case. How did the corrupt activities committed in the context of the sponsorship programme go on for such a long time without being detected or reported? It is highly unlikely that this programme could have functioned for years without any other public official involved in it not knowing about the illicit and even illegal activities committed within the administration of the programme. Why did they not denounce these activities? Were all public officials working in the programme corrupt? Finally, if public officials from inside the organization were unable or unwilling to report the corrupt activities, how did these activities escape detection by external enforcement agencies?

Although these questions are based on the specific case of the Canadian sponsorship programme scandal, they apply to many other situations in developed as well as developing countries where persistent corrupt activities in public as well as private organizations are discovered and sanctioned only after a whistle-blower reports them to the authorities and the anomalies are made public.[5]

One element to consider here is that spontaneous whistle-blowing behaviour can be difficult to initiate in practice because it implies costly activities and typically garners no economic rewards (Heyes and Kapur, 2008). The costs of whistle-blowing include monitoring and transaction costs associated with actually reporting a corrupt individual. Clearly, these costs depend on the number of agents within the organization that somehow favour the corrupt activities being committed. Indeed, the monitoring and transaction costs of reporting illicit activities can increase as more people within the organization support this type of behaviour. Thus, one way to approach this problem is to model corruption as a *social norm of behaviour*, assuming that the illicit activities are somehow accepted by most agents as valid and common practices, so that going against them means going against the majority (see, for instance, Stephan, 2008; and Davis, 1999).

[5] Recent events in Chile provide an example in the context of a developing country. At the end of 2007, widespread corruption was discovered in *Chile Deportes*, the government's sports organization. Nearly 90% of *Chile Deportes* projects had some type of anomaly (e.g., initiatives that were never started, false and nonexistent identities). It was later discovered that much of the money that went to the Valparaíso Region *Chile Deportes* branch was used to finance political campaigns of members of the Coalition of Parties for Democracy, the governing political alliance: more often known as the Concertación, this alliance has been in power since 1990. Jorge Schaulsohn, founder and former president of the Party for Democracy, was one of the first whistle-blowers to report this type of corrupt activity in the Concertación. In a press interview, Schaulsohn argued that, during the time he presided over his party, all four parties in the Concertación coalition received money from the government. Additionally, Schaulsohn reported that government money had also been used to directly finance the campaigns of Concertación candidates. He accused the government and the Concertación of an "ideology of corruption". He was later expelled from the Party for Democracy.

By analytically approaching corruption as a social norm, we can infer the following regularities. First, as seen in practice and as the economic literature on social norms recognizes, once a social norm has been established, it is very difficult to break.[6] Since most employees may be following a social norm of behaviour by not reporting corrupt activities, it can be very hard for any external enforcing agency to detect and sanction these practices. Consequently, corrupt activities can go on for many years without being detected. Second, given this inability of external enforcing agencies to detect corrupt behaviour within organizations, an inside *whistle-blower* is typically required to report these activities, thereby revealing the corrupt actions to the public. Third, since most of the employees may be complying with the established norm of behaviour, denouncing the illicit activities can be costly and typically implies punishments for whistle-blowers (e.g., ostracism; retaliation, including being fired from the organization; or even physical violence). In the case of the sponsorship scandal, Allan Cutler, the whistle-blower who lodged the complaint that triggered the departmental audit, was transferred to the technical and special services division of Public Works by the time the audit was underway and was later fired by the Canadian government.

If we assume that the whistle-blowing individuals are honest (do not accept bribes), it is clear that no payoff maximiser agent will be a whistle-blower, since this activity will only lower her payoff regardless of the whistle-blowing behaviour of the others. *This implies that, in order to study this phenomenon, a non-optimising framework is typically required.* This could explain, at least partially, the little attention that whistle-blowing behaviour has received in the theoretical economic literature on corruption.[7]

Some exceptions are the recent theoretical works by Søreide (2008) and Heyes and Kapur (2008). Søreide (2008) presents an economic framework for explaining the potential reaction of multinationals to the loss of a contract because a competitor has offered a bribe. In particular, the author examines the impact of industry structure and institutional quality on the company's incentive to react against corruption. The paper assumes a homogenous output in a standard Cournot competition framework, considering in different contexts the potential costs and benefits of a firm' whistleblowing behaviour. That study concludes, first, that firms will not react against a case of business corruption if it may disturb their opportunities to obtain cartel profits, and,

[6] See, for instance, Chapter 7 of Bowles (2004) and the many references presented therein.
[7] In a similar line, Heyes and Kapur (2008: 4) argue: "*The biggest hurdle in modeling whistle-blowing—and perhaps a reason why the phenomenon has not proven amenable to economic analysis—follows directly from…the adopted definition. If the benefits from the activity accrue, by definition, to others and not the whistle-blower, such behavior is not easily incorporated under conventional assumptions about rational, self-interested agents*."

second, the more efficient the offender of the crime, the lower the motivation for the potential whistle-blower to react. Finally, the reaction of a whistle-blower to corruption can trigger other obstacles if there are connections between local politicians and firms in the given market.

The behavioural model of Heyes and Kapur (2008) adopts, in particular, the methods of behavioural law and economics. These authors put forward a general whistle-blower "motivation function," which relates whistle-blower propensity to the characteristics of the observed malfeasance and the enforcement environment. Heyes and Kapur (2008) rely on evidence from sociology and psychology to explain why employees blow the whistle on law-breaking employers even though it is not within their narrowly defined self-interest. The authors identify three alternative "schools of thought", adapt the general motivation function to correspond to each school, and then characterize the optimal policy in each case. Their main conclusions are, first, that optimal policy varies substantially between cases and, second, the value of the information that whistle-blowers bring to the enforcement agency, and what the agency will wish to do with that information, depend upon the motives ascribed to the whistle-blowers. Finally, in adjusting the enforcement instruments, attention has to be paid to the change induced in the flow of disclosures, in addition to the direct effect on compliance incentives. In this case, the quantitative and qualitative response will also depend upon whistle-blower motives.

Our work is related to that of Heyes and Kapur (2008), as we also study the whistle-blowing behaviour of individuals within an organization but not firms' reactions, as does Søreide (2008). Nonetheless, we differ from these other articles in several respects.

First, unlike these works, we explicitly model the role and impact of whistle-blowing behaviour on the *stability of corruption*. Thus, our work is also related to those of Lui (1986), Cadot (1987), Sah (1988), Andvig and Moene (1990), Murphy et al. (1991; 1993), Acemoglu (1995), and Tirole (1996), who also emphasize the self-reinforcing nature of corruption, implying that the more people adhere to corrupt activities, the more persistent, or stable, corruption becomes. Nevertheless, including whistle-blowing behaviour in the analysis allows us to consider additional control instruments for the organization to prevent corruption, going beyond the usual variables mentioned in the economic literature (namely, the wage rate, the monitoring system or probability that an external enforcing agency will detect a corrupt individual, and penalties for corrupt activities). We attempt to answer these specific questions in this context: How can the number of whistle-blowers be increased in an organization with an initially small population of individuals who are willing to monitor and report corrupt behaviour? What additional control instruments should be considered in order to motivate and promote whistle-blowing behaviour in an organization? Furthermore, when corruption already exists as a social norm of behaviour in an organization, how can whistle-blowing policies help end the stability of corruption?

In particular, our work adds to the analysis three variables related to the whistle-blowing environment within organizations: the role of the transaction costs associated with reporting a corrupt individual, the costs of monitoring, and the economic incentives for whistle-blowers that effectively detect and sanction a corrupt agent. These economic variables can all be very important as additional control mechanisms to promote whistle-blowing behaviour and fight corruption, but their potential effect on the stability of corruption is not so clear. Thus, we wonder (and address herein): What would happen to the stability of corruption if the mechanisms used to promote whistle-blowing also incite corrupt players to "blow the whistle"? Will *corrupt whistle-blowers* contribute to break the stability of corruption or they will only add to the stability of this social norm of behaviour? How should economic instruments aimed at promoting whistle-blowing behaviour be designed so that they will effectively help break the stability of corruption rather than contribute to its stability within the organization? To the best of our knowledge, no specific work deals with the impact that economic incentives for whistle-blowers have on the stability of corruption. Consequently, this article deals with how the stability of corruption would be affected by having, or not having, these instruments. Furthermore, we analyse the effects of economic incentives on the stability of corruption considering not only *honest whistle-blowers* within an organization, but also *corrupt whistle-blowers*.[8]

Second, unlike the earlier literature, our approach to modelling social norms is based on evolutionary game theory (EGT)[9], which does not assume optimising behaviour per se, *but does retain the idea that individuals adjust their behaviour in response to persistent differentials in material incentives.* Hence, although economic agents do pursue individual material payoffs, which, in these models, represent evolutionary success, they are not always in a position to obtain straightaway the payoffs an optimising agent would obtain. This is because *social norms of behaviour* restrict the course of action of individuals in such a way as to prevent them from adjusting their behaviour towards the optimal strategy immediately (it takes time to change a social norm followed by the majority of the population).

---

[8] Heyes and Kapur (2008) also recognise this point, suggesting that a way to expand their model of whistleblower policy would be: "to explore the role played by whistle-blower regards or bounties".

[9] For further details about evolutionary game theory, see inter alia: Van Damme (1994), Vega-Redondo (1996), Weibull (1996), Samuelson (1997), and Villena and Villena (2004).

However, if this situation persists in time, some individuals will start adopting the more efficient strategy and will, therefore, receive a higher payoff than the rest of the population. In the long run, the rest of the population will start imitating this more profitable course of action. Thus, the incumbent social norm will be replaced by a new, more successful strategy that, in time, will be adopted as the new norm of behaviour in the population. This alternative EGT-based approach allows us to analyse the impact of different initial population shares ascribing to whistle-blowing behaviour within an organization and the relevance of history and initial conditions when explaining the stability of corruption.

Finally, unlike preceding works, we model corruption by assuming three basic types of behaviour within an organization: (i) *honest behaviour*, which implies that an individual does not receive any bribe from illicit activity; (ii) *corrupt behaviour*, in which case an individual does receive a bribe from a corrupt relationship; and (iii) *whistle-blowing (honest-enforcer) behaviour*, with an agent who not only behaves honestly (i.e., does not receive bribes), but also monitors other people within the organization and reports them if they are behaving corruptly. Assuming that an organization includes three different population shares pursuing these three alternative behaviour types, we use the concept of replicator dynamics to analyse which population share will become stable within the organization in the long run. In other words, we formally explore the asymptotical stability of *the non-corruption equilibrium* (all individuals behave honestly) and the *corruption equilibrium* (all individuals behave corruptly). We also extend this basic model to include cases in which corrupt individuals also ascribe to whistle-blowing behaviour. In order to model this type of behaviour, we modify the previous strategic setting by assuming that the organization encompasses a proportion of *honest individuals*, another of *corrupt whistle-blowing individuals*, and another of *honest whistle-blowing individuals*.

Section 2 of this paper describes the basic economic model and its results are given in Section 3. In Section 4, we set forward an extension of the basic model, and Section 5 presents some concluding remarks and related topics for future research.

2. The Model

While the model we present in this work is quite general and can be applied to several contexts, in order to ease the exposition henceforth we consider the case of a public organization. We assume that there exists an infinitely elastic demand for corrupt services and that individuals or firms interested in buying these services can buy them at a price of $\beta > 0$, the value of the bribe paid to a public servant. Clearly, for these bribes to be paid by rational individuals, the expected benefits obtained from corrupt services must be higher than the expected costs. In this context, we can think of corrupt services as the action, vote, or influence of a person in an official or public capacity in order to, for instance, bypass laws and regulations, obtain government contracts, acquire state-owned property, etc. In any case, the cost of the bribery should be lower than the costs associated with obtaining the services lawfully; otherwise rational individuals would not buy corrupt services.

On the supply side, we assume that there exist a number of bureaucrats that may offer corrupt services. Suppose that, whereas honest bureaucrats earn the wage w, corrupt bureaucrats earn the same wage w plus the amount of the bribery, $\beta > 0$. Now, let us assume that the government, through an external enforcing agency, for example, can detect a corrupt public official within a specific governmental organization with probability θ. Whenever the bureaucrat is detected by the government, she will be fired and have to go to work in the private sector, where she can get a wage $w_0 \geq 0$. In addition, the corrupt bureaucrat will also have to pay a penalty $f \geq 0$.

Clearly, in many real life situations, as those discussed in the introduction, it can be very hard for the government to detect corrupt government officials. In terms of our model, we can argue that there are typically a great number of bureaucrats; even more firms being regulated, applying for government contracts, bidding on state-owned property, etc.; and limited resources for monitoring and auditing both public officials and firms. But even if the government had a great deal of resources with which to investigate, the important informational advantage of public officials can make it too expensive to completely eliminate corruption from the system. Indeed, designing an effective control and monitoring system can be both difficult and costly, which may imply a very low θ. Here, given these contexts, the role of whistle-blowing can be very important.

Assume, for instance, that there are some whistle-blowing bureaucrats that monitor others and denounce them to the government if found in an illicit activity, for instance providing corrupt services to private firms. If a whistle-blowing bureaucrat can detect a corrupt public official within her own organization with probability Θ, this probability being greater than the probability of detection by an external enforcing agency, that is $\Theta > \theta$, then the expected gains from corrupt activities decrease in comparison with the baseline scenario of no whistle-blowing bureaucrats in the organization.

Nevertheless, it is difficult for whistle-blowing behaviour to arise spontaneously. As we already discussed in the introduction, this is because the whistle-blowing strategy is a costly activity. The costs associated with whistle-blowing behaviour include monitoring costs, $m \geq 0$, and the transaction costs associated with actually reporting a public official to the government, $\tau > 0$. If we assume that the whistle-blowing bureaucrats are honest, in the sense that they do not accept bribes from private firms for instance, it is clear that no payoff maximiser agent will be a whistle-blower, since this activity will only lower her payoff regardless of the whistle-blowing behaviour of the rest. In fact, in this case, the whistle-blowing strategy is strictly dominated by the honest strategy $(w > w - \tau\Theta - m)$, so no honest bureaucrat will have incentives to become a whistle-blower. This, in turn, implies that no bureaucrat will be deterred from corruption by the threat of whistle-blowing behaviour from her colleagues.

2.1 The Evolutionary Game Population Dynamics

We now adopt an approach based on an evolutionary game theoretic framework that allows us to model corruption as a social norm of behaviour. In terms of the specific strategies pursued by the bureaucrats, in this case, norms of behaviour, we suppose that within the government there is a proportion of *honest bureaucrats*, another of *corrupt bureaucrats*, and another of *whistle-blowing* (*honest-enforcer) bureaucrats*, which we denote p_1, p_2, and p_3, respectively. Considering a total population of n public officials and given the population shares p_i at any point in time, it is assumed that each of the $p_1 n$ honest bureaucrats receives the payoff associated with honest behaviour, namely the wage w.

By contrast, the $p_2 n$ corrupt bureaucrats receive the expected payoff associated with corrupt behaviour, namely $(1-\theta)(w+\beta)+(\theta)(w_0 - f)$. In addition, they also perceive the expected costs of being detected by the whistle-blowing bureaucrats and reported to the government. The corrupt bureaucrats are monitored by the $p_3 n$ whistle-blowing bureaucrats, and so they can be sanctioned by the government if one of these agents catches the bureaucrat in a corrupt relationship. Herein, we assume that a corrupt bureaucrat will meet a whistle-blowing bureaucrat with probability p_3 and that the latter will detect and report a corrupt agent with probability Θ. Hence, the net expected payoff associated with the corrupt strategy becomes $\pi_2 = (1-\theta - p_3\Theta)(w+\beta) + (\theta + p_3\Theta)(w_o - f)$.

Finally, each of the $p_3 n$ whistle-blowing bureaucrats receives the payoff associated with honest behaviour, namely wage w (we relax this assumption later). The costs associated with monitoring and reporting a corrupt bureaucrat are the transaction costs associated with reporting a colleague to the government, $\tau > 0$, and the monitoring costs, denoted by $m > 0$, which we assume in this context to be an increasing function of the proportion of corrupt bureaucrats, p_2, that is $m(p_2) \geq 0$, with $\frac{dm(p_2)}{dp_2} > 0$ and $m(0) = 0$.[10] This last assumption is consistent with previous works on corruption that suppose that it is harder to audit corrupt officials in societies where corruption is more prevalent (Lui, 1986; Cadot, 1987; Andvig and Moene, 1990).

We also include here a reward for the whistle-blowing bureaucrats subject to the effective detection and sanction of a corrupt agent, denoted by $\sigma \geq 0$. This reward can be thought of as any incentive given by the government to bureaucrats for monitoring the work done by their colleagues and reporting illicit acts. We analyse what happens in terms of the stability of the corruption and non-corruption equilibria if this mechanism is in place and if it is not. Hence, if a whistle-blowing bureaucrat will meet a corrupt agent with probability p_2, and will detect and report her with probability Θ, the net expected gains for whistle-blowing bureaucrats are: $\pi_3 = w + p_2\Theta(\sigma - \tau) - m(p_2)$. The payoffs for each strategy type, given the population composition, are therefore:

(1) $\pi_1 = w$

(2) $\pi_2 = (1-\theta - p_3\Theta)(w+\beta) + (\theta + p_3\Theta)(w_o - f)$

(3) $\pi_3 = w + p_2\Theta(\sigma - \tau) - m(p_2)$

[10] We assume that to actually report a corrupt bureaucrat, individuals who choose the whistle blowing strategy must develop monitoring activities. However, monitoring is costly. Agents that monitor must spend time, for instance, investigating and proving suspicious behaviour, accusations, etc. and might also face other costs like verbal violence and threats from the group of corrupt agents. These costs are likely to increase along with the fraction of the population that is behaving corruptly.

Equations (1), (2), and (3) clearly show that, if there are no corrupt bureaucrats $(p_2 = 0)$, honest bureaucrats will perform as well as whistle-blowing bureaucrats. By contrast, if corrupt bureaucrats are present in the population, the whistle-blowing strategy is weakly dominated by the honest strategy whenever the transaction costs associated with reporting a public official to the government are larger or equal to the monetary reward for detecting and sanctioning a corrupt bureaucrat, that is $\sigma \leq \tau$.

Let us now formalise the replicator equation as typically presented in the evolutionary game theoretical literature[11]. Consider an evolutionary game with *n* pure strategies and stage game pay-off π_{ij} for any *i*-player who meets any *j*-player. If $p = (p_1, \ldots, p_n)$ is the frequency of each type in the population, the expected payoff for the *i*-player is then $\pi_i(p) = \sum_{j=1}^{n} p_j \pi_{ij}$, and the average payoff in the game is $\overline{\pi}(p) = \sum_{i=1}^{n} p_j \pi_i(p)$. The replicator dynamic for this game is then given by:

$$(4) \qquad \dot{p}_i = p_i\left(\pi_i(p) - \overline{\pi}(p)\right)$$

The replicator equation expresses the idea that strategies grow in the population if they do better than average; the strategies that do best grow fastest. One immediately sees that a Nash equilibrium is a stationary point in the dynamic system. Conversely, each stable stationary point is a Nash equilibrium and an asymptotically stable fixed point is a perfect equilibrium.

In the context of equations (1), (2), and (3), the replicator dynamics will be represented by the two differential equations presented below:

$$(5) \quad \frac{dp_1}{dt} = \dot{p}_i = p_1\left(\pi_1 - \overline{\pi}\right)$$

$$(6) \quad \frac{dp_2}{dt} = \dot{p}_2 = p_2\left(\pi_2 - \overline{\pi}\right)$$

[11] The mathematical formulation of the replicator dynamics is due to Taylor and Jonker (1978). For details, see also, Vega-Redondo (1996), Weibull (1996), and Gintis (2000).

where $\overline{\pi} = p_1\pi_1 + p_2\pi_2 + (1 - p_1 - p_2)\pi_3$ is the average payoff in the population as a whole[12]. From (5) and (6), it is clear that the rate of growth of the share of the population using strategies 1 and 2 are proportional to the amount by which those strategies' payoffs exceed the average payoff of the strategies in the population.

3. Results

Taken together, equations (5) and (6) constitute *a system of first order differential equations*.[13] From these equations we have that: $\frac{\partial \dot{p}_1}{\partial p_1} = \pi_1 - \overline{\pi} + p_1 \frac{\partial(\pi_1 - \overline{\pi})}{\partial p_1}$; $\frac{\partial \dot{p}_1}{\partial p_2} = p_1 \frac{\partial(\pi_1 - \overline{\pi})}{\partial p_2}$; $\frac{\partial \dot{p}_2}{\partial p_1} = p_2 \frac{\partial(\pi_2 - \overline{\pi})}{\partial p_1}$; and $\frac{\partial \dot{p}_2}{\partial p_2} = \pi_2 - \overline{\pi} + p_2 \frac{\partial(\pi_2 - \overline{\pi})}{\partial p_2}$. Hence, from equations (1), (2), and (3) and the definition of $\overline{\pi}$, the *Jacobian* of $\left(\dot{p}_1, \dot{p}_2\right)$ at an arbitrary point $\left(p_1^*, p_2^*\right)$ is:

(7) $$J\left(\dot{p}_1, \dot{p}_2\right)\left(p_1^*, p_2^*\right) = \begin{pmatrix} \pi_1 - \overline{\pi} + p_1(-\pi_1 + \pi_3) & p_1(-\pi_2 + \pi_3) \\ p_2(-\pi_1 + \pi_3) & \pi_2 - \overline{\pi} + p_2(-\pi_2 + \pi_3) \end{pmatrix}$$

In particular, we want to examine the stability of two equilibria: the case in which no corrupt agents are present, that is $(p_1 = a, p_2 = 0)$ with $a \in [0,1]$, and the case in which the population consists only of corrupt bureaucrats, that is $(p_1 = 0, p_2 = 1)$.[14] The stability of these two equilibria is examined in the following sections.

---

[12] It can be easily verified that the population shares p_i always add up to one and remain nonnegative under the replicator dynamics.

[13] It is "first order" because no derivatives higher than the first appear. It is "ordinary" as opposed to "partial" because we want to solve for a function of the single variable *t*, as opposed to solving for a function of several variables.

[14] The only remaining candidate for a stable equilibrium is $(p_1 = 0, p_2 > 0)$, consisting exclusively of corrupt and honest-enforcer players. We do not address this case here.

3.1 The Stability of the Non-corruption Equilibrium

Proposition 1: *The non-corruption equilibrium* $\left(p_1 = a, p_2 = 0\right)$ *with* $a \in \left[0,1\right]$*, i.e. no corrupt agents are present in the population, is local asymptotically stable if and only if:*

(8) $\left(1-\theta-p_3\Theta\right)\left(w+\beta\right)+\left(\theta+p_3\Theta\right)\left(w_o-f\right)<w$

***Proof of Proposition 1*:** The steady stable equilibrium $\left(p_1 = a, p_2 = 0\right)$ with $a \in \left[0,1\right]$, implies that $\bar{\pi} = \pi_1 = \pi_3$. In this case, the *Jacobian* shown in (7) becomes $J\left(\dot{p}_1, \dot{p}_2\right)\left(p_1^*, p_2^*\right) = \begin{pmatrix} -a\left(\pi_1-\pi_3\right) & -a\left(\pi_2-\pi_3\right) \\ 0 & \pi_2-\pi_1 \end{pmatrix} = \begin{pmatrix} 0 & a\left(\pi_3-\pi_2\right) \\ 0 & \pi_2-\pi_1 \end{pmatrix}$. Hence, the determinant of the *Jacobian* is zero, which implies that one eigenvalue is zero and the other equals the trace of the *Jacobian*. Thus, the inequality $\pi_2 < \pi_1$ is necessary and sufficient for local asymptotic stability, which, in turn, implies the condition in (8). **Q.E.D.**

Proposition 1 shows that the non-corruption equilibrium will be local asymptotically stable whenever the expected net benefits from behaving corruptly in a population that consists only of honest and whistle-blowing agents, i.e. $\left(1-\theta-p_3\Theta\right)\left(w+\beta\right)+\left(\theta+p_3\Theta\right)\left(w_o-f\right)$, are lower than the benefits from behaving honestly, i.e. the government salary, w. From this equilibrium condition, it can inferred that efficiency wages for government officials, β, a high probability of being detected by the government, θ, and a high penalty for corrupt public officials, f, can imply significant expected costs of dismissal and so that the non-corruption equilibrium will be fairly stable and hard to break. By contrast, whenever the amount of the bribery, β, and the salary the government official can get from the private sector if dismissed, w_0, are high in comparison with the salary paid by the government and the penalty for being caught in a corrupt activity the non-corruption equilibrium will not be very stable.

In terms of the effect of whistle-blowing behaviour on the local asymptotically stability of the non-corruption equilibrium, it can be noted that in this case, the corrupt player confronts two expected costs: the expected costs from being detected by an external enforcing governmental agency, $\theta\left(w_0 - f\right)$, and the expected costs from being detected and reported to the government by a whistle-blower from within the organization, $p_3\Theta\left(w_0 - f\right)$. From this latter type of cost, typically absent in the traditional economic literature on corruption,[15] it is clear that the stability of the non-corruption equilibrium crucially depends upon the *population share ascribing to the whistle-blowing strategy*, which allows us to highlight some basic remarks on the economics of whistle-blowing .

Government organisations with a large population of honest bureaucrats that not only do not accept bribes but that also are willing to monitor and report corrupt behaviour ensure higher expected costs for corrupt agents than the baseline case in which whistle-blowing is not a common practice. In this context, given a low probability of being detected by an external enforcing governmental agency and, therefore, low associated expected costs, the non-corruption equilibrium can still be local asymptotically stable due to the expected costs inflicted on corrupt agents by whistle-blowers. Indeed, if $\theta \to 0$, the condition for the local asymptotic stability of the non-corruption equilibrium becomes: $\left(1 - p_3\Theta\right)\left(w + \beta\right) + \left(p_3\Theta\right)\left(w_o - f\right) < w$. From this condition is clear that provided that the whistle-blowing population within the organization, p_3, is high enough to ensure that the expected costs to corrupt agents can still be higher than the expected net benefits from behaving corruptly in a population that consists only of honest agents, the non-corruption equilibrium can still be very stable. This implies that whenever honesty and whistle-blowing behaviour becomes a common practice, i.e. becomes a social norm; it can be very difficult to stop, even despite low probabilities of detection by external enforcing agencies.

As the design of a successful external control and monitoring mechanism can be both difficult and costly, and in practice may imply a very low θ, it can be argued that the expected costs for corrupt agents resulting from being detected and reported to the government by a whistle-blower are always higher than the expected costs of being detected by an external enforcing governmental agency, that is $\theta\left(w_0 - f\right) < p_3\Theta\left(w_0 - f\right)$. This implies that whistle-blowing behaviour can indeed in some contexts be more efficient for deterring corruption than the typical government monitoring system.

Finally, proposition 1 also points to the relevance of history and initial conditions when explaining the stability of the corruption or non-corruption equilibria. Indeed, as suggested in

[15] See for instance the recent survey on the economics of corruption by Aidt (2003).

previous works, such as those of Acemoglu (1995) and Tirole (1996), in this model, history is understood as the past behaviour of the member's group and is an important determinant of the group's current behaviour. Specifically, from condition (8) we can infer that a low initial population of whistle-blowers may imply that the non-corruption equilibrium, in which no corrupt agents are present in the population, may never be stable in some contexts. Why is it that this population of honest bureaucrats (who not only do not accept bribes but also are willing to monitor and report corrupt behaviour) can be initially small in some organizations and societies and very significant in others? What factors determine these diverse norms of behaviour in different organizations and societies? Clearly, to answer these questions, we have to consider the historical background in the analysis. What is clear, though, is that the current state of affairs must be necessarily analysed considering related events that occurred in the past and that somehow determine the present situation. Here, initial conditions are crucial and even small differences may imply widely differing outcomes.

3.2 The Stability of the Corruption Equilibrium

Proposition 2: *The corruption equilibrium, $p_1 = 0$, i.e. all the population behaving corruptly, is local asymptotically stable if and only if:*

i) $w < (1-\theta)(w+\beta)+(\theta)(w_0 - f)$ whenever $\Theta(\sigma-\tau)-m(1)\leq 0$, and

ii) $w+\Theta(\sigma-\tau)-m(1) < (1-\theta)(w+\beta)+(\theta)(w_0 - f)$ whenever $\Theta(\sigma-\tau)-m(1) > 0$.

***Proof of Proposition 2*:** The steady stable equilibrium $(p_1 = 0, p_2 = 1)$ implies that $p_3 = 0$ and $\bar{\pi} = \pi_2$. In this case, the *Jacobian* shown in (7) becomes $J(\dot{p}_1, \dot{p}_2)(p_1^*, p_2^*) = \begin{pmatrix} \pi_1 - \pi_2 & 0 \\ \pi_3 - \pi_1 & \pi_3 - \pi_2 \end{pmatrix}$. The conditions for the *Jacobian* matrix to have a positive determinant and a negative trace are: (i) $\pi_1 - \pi_2 < 0$ and (ii) $\pi_3 - \pi_2 < 0$. The former implies that $w < (1-\theta)(w+\beta)+(\theta + p_3\Theta)(w_0 - f)$, as $p_3 = 0$ we have that:

(9) $w < (1-\theta)(w+\beta)+(\theta)(w_0 - f)$

The latter implies that $w + p_2\Theta(\sigma - \tau) - m(p_2) < (1-\theta)(w+\beta)+(\theta + p_3\Theta)(w_0 - f)$, as $p_2 = 1$ and $p_3 = 0$, we obtain:

(10) $w + \Theta(\sigma - \tau) - m(1) < (1-\theta)(w+\beta)+(\theta)(w_0 - f)$

Hence, we have two different cases to consider:

Case i). If $\Theta(\sigma - \tau) - m(1) > 0$, then equation (10) is the only necessary and sufficient condition for local asymptotic stability.
Case ii). If $\Theta(\sigma - \tau) - m(1) \leq 0$, then equation (9) is the only necessary and sufficient condition for local asymptotic stability **Q.E.D.**

Corollary 1: *In the absence of any incentives to whistle-blowing behaviour,* $\sigma = 0$*, the corruption equilibrium,* $p_1 = 0$*, i.e. all the population behaving corruptly, is local asymptotically stable if and only if:* $w < (1-\theta)(w+\beta)+(\theta)(w_0 - f)$.

Proof of Corollary 1: Whenever $\sigma = 0$, the inequality $-\Theta\tau - m(1) < 0$ is always satisfied. From proposition 2, this implies that $w < (1-\theta)(w+\beta)+(\theta)(w_0 - f)$ is the only condition for the corruption equilibrium to be local asymptotically stable **Q.E.D.**

From condition (i) in proposition 2, it is clear that the local asymptotical stability of the corruption equilibrium is contingent on the environment for whistle-blowing behaviour. If the costs associated with whistle-blowing are larger than or equal to the benefits, the corruption equilibrium can be very stable. Indeed, if the costs of monitoring a population of only corrupt agents and the transaction costs associated with actually reporting a corrupt government official, $\Theta\tau + m(1)$, are high in comparison with the expected benefits of whistle-blowing behaviour, $\Theta\sigma$, then the condition for corruption stability will become: $w < (1-\theta)(w+\beta)+(\theta)(w_0 - f)$. Corollary 1 also shows that, in the absence of incentives for whistle-blowers, $\sigma = 0$, this inequality will be the only necessary and sufficient condition for local asymptotic stability.

From this condition, it becomes clear that, if the probability of being detected by the government is rather small, the corruption equilibrium will be fairly stable and hard to break. In fact, if $\theta \to 0$, the condition for the local asymptotic stability of the corruption equilibrium is always satisfied: $w < w + \beta$, which implies that whenever corruption becomes a common practice, it will be very difficult to stop. In the same way, if the amount of the bribery, β, and the salary the government official can get from the private sector if dismissed, w_0, are high in comparison with the salary paid by the government, w, and the penalty for being caught in a corrupt activity, f, the corruption equilibrium will also be very stable.

By contrast, increased salaries for government officials may break the stability of corruption. This argument dates back to Becker and Stigler (1974), who pointed out that efficiency wages can be used to control corruption since they increase the cost of dismissal and, therefore, make bureaucrats more reluctant to accept bribes[16]. Nevertheless, paying efficiency wages can be very expensive for governments and does not ensure that corruption will be reduced in all situations (see, for instance, Mookherjee and Png, 1995). In addition, if the probability of being detected by the government is rather small, the potential effect of efficiency wages on the stability of corruption is reduced, since the expected cost of dismissal also becomes very small.

Likewise, an increase in the penalty for corrupt public officials can also break the stability of corruption. An interesting point to notice here is that if legal punishments for corrupt officials are high enough, then efficiency wages are not needed to reduce corrupt behaviour and break its stability as a norm of behaviour. Indeed, assuming that the penalty for corrupt behaviour is equal to the bribe, $f = \beta$, and that the government only pays reservation wages, $w = w_0$, corruption will never be stable as long as the probability of being detected by the government is greater than or equal to 50%, that is $\theta \geq \frac{1}{2}$. Here again, if the ability of the government to detect

[16] For favourable empirical evidence of this point see, for instance, van Rijckeghem and Weder (2001).

corrupt officials is reduced, the impact of penalties and legal punishments in general in the stability of corruption will be quite limited as well.

Again, as the design of a successful control and monitoring mechanism can be both difficult and costly, and in practice may imply a very low θ, this result helps explain why corruption, once it has become a common practice in a government, is very hard to stop and can even last for many years.

However, key to this result was the assumption of an unfavourable environment for whistle-blowing behaviour, namely $\Theta(\sigma - \tau) - m(1) \leq 0$. What happens if this is not the case? Condition (ii) in proposition 2 shows that the corruption equilibrium will be local asymptotically stable whenever the expected net benefits from behaving corruptly in a population that consists only of corrupt players, i.e. $(1-\theta)(w+\beta)+(\theta)(w_0 - f)$, are greater than the expected net benefits from behaving honestly and performing whistle-blowing activities, i.e. the government's salary w and $\Theta(\sigma - \tau) - m(1)$.

This condition for the stability of corruption brings several new elements to the discussion. First, there are additional control instruments for the government to consider apart from the wage rate (w),monitoring system (θ), and penalty (f) discussed above. These include the transaction costs associated with reporting a corrupt public official to the government, $\tau > 0$, monitoring costs, $m \geq 0$, and incentives for whistle-blowers that effectively detect and sanction a corrupt agent, $\sigma \geq 0$.

These alternative control instruments for the government imply new challenges in terms of policy. Here, for instance, it could be argued that all the countries that have formally established *legal protection for whistle-blowers* have reduced the transaction costs for whistle-blowers to report corruption, which implies a lower τ. Similarly, *more government transparency and freedom of information legislation* points to reduced monitoring costs, which imply a lower m. Nevertheless, what is typically not seen as in policies against the establishment of corruption as a norm of behaviour in government are incentives for whistle-blowing behaviour. As shown in our model, this is clearly beyond legal protection for whistle-blowers, which only reduces the transaction costs for whistle-blowers to report corruption, indicating that whistle-blowing is not a bad thing but, on the contrary, something to reward. Hence, the first thing to consider is the common reaction to whistle-blowing. Whereas some people see whistle-blowers as selfless martyrs for public interest and organizational accountability, others view them as "snitches", solely pursuing personal glory and fame. Incentives to whistle-blowing behaviour should start with the government's public recognition that reporting corruption is positive and should be encouraged at all levels. This message should be delivered not only to the people working within government institutions but also to the public in general. In addition to this, economic incentives should also play a part in encouraging whistle-blowing behaviour.[17] However, this is a complex economic policy issue that, as we will see in the next section, can be difficult to design and apply, and if poorly implemented can contribute to the stability of corruption instead of its eradication.

Second, from condition (ii) in proposition 2, we can also infer that if a whistle-blower can detect a corrupt public official with a probability greater than that of detection by an external agency from the government, that is $\Theta > \theta$, then the effect of whistle-blowing on the stability of corruption could be quite significant, especially when considering that the more traditional control instruments available to governments for fighting corruption are extremely dependent of the ability of the government to detect corrupt officials. Clearly, a member inside the organization is more likely to spot corruption, and spot it at an earlier stage, than an external enforcing agency from the government.

Third, whistle-blowing can certainly be cheaper than designing a successful control and monitoring mechanism or paying efficiency wages. In fact, if, for instance, the fines charged to corrupt public officials are paid as incentives to the whistle-blowers that detect them, that is $\sigma = f$, then a monitoring system based on whistle-blowing can be both very effective and low cost for the government.

Finally, as typically recognised in the literature, sometimes a "big push" is required in order to reduce the level of corruption in societies where it is epidemic (for details and further references see, for instance, Aidt, 2003). In the context of our model, this "big push" needed to break the stability of corruption can come from encouraging whistle-blowing behaviour in government institutions. Indeed, from condition (ii) in proposition 2, it follows that, given sufficiently large expected benefits from effectively reporting a corrupt agent, that is, the net value of the transaction and monitoring costs, $\Theta\left(\sigma-\tau\right)-m\left(1\right)$, the stability of corruption where

[17] Although this is not yet common practice, there are currently some examples, such as the HOPE Scholarship in Georgia, which provides four years of free tuition to a tech school or University in Georgia for children of whistleblowers or those researching government, corporate, or religious crimes. For details see: http://www.gsfc.org/hope/.

all the population behaves corruptly can be broken and, therefore, corruption will no longer be the social norm of behaviour within the organization.

4. An Extension of the Model

We now extend the model presented in the previous section by considering the case in which *corrupt bureaucrats* also ascribe to whistle-blowing behaviour. In order to model this type of behaviour, we modify the previous strategic setting by assuming that, within the government, there is a proportion of *honest bureaucrats* (p_1), another of *corrupt whistle-blowing bureaucrats* (p_2), and another of *honest whistle-blowing bureaucrats* (p_3). Consequently, we assume that the whistle-blowing bureaucrats can be either corrupt or honest (they may or may not accept bribes).

The payoffs for each strategy type, given the new population composition, are:

(11) $\pi_1 = w$

(12) $\pi_2 = (1-\theta - p_3\Theta)(w+\beta) + (\theta + p_3\Theta)(w_o - f) + p_2\Theta(\sigma - \tau) - m(p_2)$

(13) $\pi_3 = w + p_2\Theta(\sigma - \tau) - m(p_2)$

From equations (11), (12), and (13), the *Jacobian* of $\left(\dot{p}_1, \dot{p}_2\right)$ at an arbitrary point $\left(p_1^*, p_2^*\right)$ now becomes:

$$(14)\ J\left(\dot{p}_1,\dot{p}_2\right)\left(p_1^*,p_2^*\right) = \begin{pmatrix} \pi_1 - \bar{\pi} + p_1(-\pi_1 + \pi_3) & p_1(-\pi_2 + \pi_3) \\ p_2(\pi_3 - \pi_1) & \pi_2 - \bar{\pi} + \left(\Theta(\sigma - \tau) - \dfrac{\partial m(p_2)}{\partial p_2} + \pi_3 - \pi_2\right) \end{pmatrix}$$

It can be easily shown that the basic condition for the local stability of the non-corruption equilibrium, in which no corrupt agents are present, i.e., $(p_1 = a, p_2 = 0)$, with $a \in [0,1]$, remains the same as before; that is $(1-\theta - p_3\Theta)(w+\beta) + (\theta + p_3\Theta)(w_o - f) < w$, see proposition 1. Hence, we will only focus on the condition for the local stability of the corruption equilibrium, or the case in which the population consists only of corrupt bureaucrats, i.e., $(p_1 = 0, p_2 = 1)$.

Proposition 3: *The corruption equilibrium* $(p_1 = 0, p_2 = 1)$*, i.e., all the population behaving corruptly, is local asymptotically stable if and only if:*

i) $w + \Theta(\sigma - \tau) - \frac{\partial m(1)}{\partial p_2} < (1-\theta)(w+\beta) + (\theta)(w_o - f)$ whenever $\Theta(\sigma - \tau) = m(1) > \frac{\partial m(1)}{\partial p_2}$

and,

ii) $w < (1-\theta)(w+\beta) + (\theta)(w_o - f) + \Theta(\sigma - \tau) - m(1)$ whenever $\Theta(\sigma - \tau) = \frac{\partial m(1)}{\partial p_2} > m(1)$.

***Proof of Proposition 3*:** The steady stable equilibrium $(p_1 = 0, p_2 = 1)$ implies that $p_3 = 0$ and $\bar{\pi} = \pi_2$. In this case, the *Jacobian* shown in (14) becomes $J\left(\dot{p}_1, \dot{p}_2\right)\left(p_1^*, p_2^*\right) = \begin{pmatrix} \pi_1 - \pi_2 & 0 \\ \pi_3 - \pi_1 & \Theta(\sigma - \tau) - \frac{\partial m(p_2)}{\partial p_2} + \pi_3 - \pi_2 \end{pmatrix}$. The conditions for the *Jacobian* matrix to have a positive determinant and a negative trace are: (i) $\pi_1 - \pi_2 < 0$ and (ii) $\Theta(\sigma - \tau) - \frac{\partial m(p_2)}{\partial p_2} + \pi_3 - \pi_2 < 0$. The former implies that $w < (1 - \theta - p_3\Theta)(w + \beta) + (\theta + p_3\Theta)(w_o - f) + p_2\Theta(\sigma - \tau) - m(p_2)$, as $p_2 = 1$ and $p_3 = 0$, we have that:

(15) $w < (1-\theta)(w+\beta) + (\theta)(w_o - f) + \Theta(\sigma - \tau) - m(1)$

The latter implies that $w + \Theta(\sigma - \tau) - \frac{\partial m(p_2)}{\partial p_2} < (1 - \theta - p_3\Theta)(w + \beta) + (\theta + p_3\Theta)(w_o - f)$, as $p_2 = 1$ and $p_3 = 0$, we obtain:

(16) $w+\Theta(\sigma-\tau)-\frac{\partial m(1)}{\partial p_2}<(1-\theta)(w+\beta)+(\theta)(w_o-f)$

Hence, we have two different cases to consider:

Case i). If $\Theta(\sigma-\tau)=m(1)>\frac{\partial m(1)}{\partial p_2}$, then equation (16) is the only necessary and sufficient condition for local asymptotic stability.

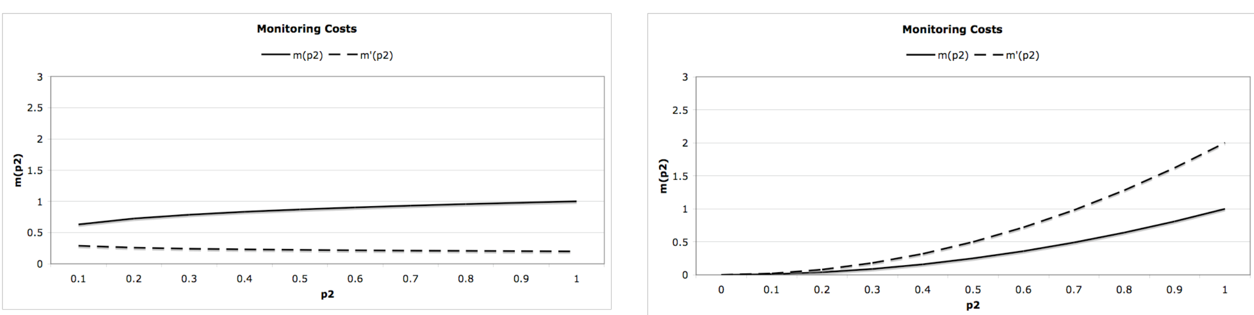
Case ii). If $\Theta(\sigma-\tau)=\frac{\partial m(1)}{\partial p_2}>m(1)$, then equation (15) is the only necessary and sufficient condition for local asymptotic stability **Q.E.D.**

From conditions (i) and (ii) in proposition 3, we can see that the effect that the mechanisms used to encourage whistle-blowing have on the local asymptotic stability of corruption depends on the nature of the monitoring costs within the organization. Whenever condition $m(p_2)>\frac{\partial m(p_2)}{\partial p_2}$ is satisfied, i.e., when monitoring costs exceed marginal monitoring costs, marginal monitoring costs are a diminishing function of the corrupt bureaucrat population; see Figure 1 (panel a) below.[18] As marginal monitoring costs are bounded from above, the expected net benefits from detecting and reporting a corrupt bureaucrat do not need to be set very high in order to ensure that condition $\Theta(\sigma-\tau)=m(p_2)$ is satisfied. At the limit, where all the population behaves corruptly ($p_2=1$) the condition $\Theta(\sigma-\tau)=m(1)>\frac{\partial m(1)}{\partial p_2}$ must hold. In this case, the economic incentives for whistle-blowers can be rather low and, hence, the level of corrupt bureaucrats ascribing to whistle-blowing behaviour can also be limited. This, according to condition (i), make the corruption equilibrium less stable, as can be easily inferred from the condition $w+\Theta(\sigma-\tau)-\frac{\partial m(1)}{\partial p_2}<(1-\theta)(w+\beta)+(\theta)(w_o-f)$, where the expected net benefits from

[18] In particular, in this case the monitoring costs function could be formally defined as follows: $m(0)=0;\frac{\partial m(p_2)}{\partial p_2}>0;\frac{\partial^2 m(p_2)}{\partial p_2^{\,2}}<0;\quad and\quad Lim_{p_2\to\infty}\frac{\partial m(p_2)}{\partial p_2}=0$.

being a whistle-blower in a population of only corrupt agents, i.e., $\Theta(\sigma-\tau)-\dfrac{\partial m(1)}{\partial p_2}$, makes corrupt behaviour less attractive in economic terms and, therefore, the *corruption equilibrium* becomes less stable.

By contrast, if condition $\dfrac{\partial m(p_2)}{\partial p_2} > m(p_2)$ is satisfied, that is, when monitoring marginal costs lie above monitoring costs, monitoring costs increase along with the population of corrupt bureaucrats, not being bounded from above. In other words, marginal monitoring costs are an increasing function of the corrupt bureaucrat population; see Figure 1 (panel b). In this case, in order to make whistle-blowing behaviour economically "attractive", the expected net benefits of detecting and reporting a corrupt bureaucrat, $\Theta(\sigma-\tau)$, may have to be set very high. At the limit where all the population behaves corruptly, $p_2 = 1$, condition $\Theta(\sigma-\tau)=\dfrac{\partial m(1)}{\partial p_2} > m(1)$ must hold. In this case, economic incentives for whistle-blowers may have to be set so high that many corrupt bureaucrats will ascribe to whistle-blowing behaviour, which, according to condition (ii), makes the corruption equilibrium more stable. This can be noted from condition $w < (1-\theta)(w+\beta)+(\theta)(w_o - f)+\Theta(\sigma-\tau)-m(1)$, where the expected net benefits from being a whistle-blower in a population of only corrupt agents, i.e., $\Theta(\sigma-\tau)-m(1)$, make corrupt behaviour more attractive in economic terms and, therefore, the corruption equilibrium becomes more stable.



(a) **(b)**

Figure 1: Examples of Monitoring Cost Functions

From proposition 3, it is clear that there are cases in which the governmental control instruments aimed at motivating whistle-blowing behaviour can be counter-productive, making corruption more stable within an organization. In particular, a high level of economic incentives for whistle-blowers can increasingly induce corrupt agents to become what we have called *corrupt whistle-blowers*; this, in turn, can make honest behaviour and honest whistle-blowing behaviour less economically attractive.

According to proposition 3, in order to know in what contexts *economic incentives for whistle-blowers* can contribute more effectively to making corruption less stable within an organization, it is vital to understand *the nature of monitoring costs*. In particular, in contexts in which marginal monitoring costs are a diminishing function of the corrupt bureaucrat population, economic incentives for whistle-blowers can make corrupt behaviour less economically attractive and, therefore, the corruption equilibrium becomes less stable in the organization. By contrast, whenever marginal monitoring costs are an increasing function of the corrupt bureaucrat population, economic incentives for whistle-blowers can make corrupt behaviour more attractive in economic terms and, therefore, the corruption equilibrium in the organization becomes more stable. Consequently, a policy recommendation in order to adequately design economic incentives for whistle-blowers would be to investigate monitoring costs in different settings in detail. Clearly, this is a matter to be resolved empirically.

5. Concluding Remarks

This paper developed a simple evolutionary game theoretic framework to investigate the role and potential impact of whistle-blowing on the persistence of corruption in organizations. We can draw several conclusions based on the results found herein.

First, it is difficult for whistle-blowing behaviour to arise spontaneously in practice because this strategy is a costly activity and typically involves no economic rewards. The costs of whistle-blowing include the monitoring and transaction costs associated with actually reporting a public official to the government. If we assume that the whistle-blowing bureaucrats are honest (do not accept bribes), it is clear that no payoff maximiser agent will be a whistle-blower, since this activity will only lower her payoff regardless of the whistle-blowing behaviour of the others. *This implies that, in order to study this phenomenon, a non-optimising framework is required. Thus, we modelled corruption as a social norm of behaviour using an evolutionary game theoretic framework.*

Second, we model corruption by assuming three basic types of behaviour within an organization: (i) honest behaviour; (ii) corrupt behaviour; and (iii) honest whistle-blowing behaviour. From this analysis, we conclude that the non-corruption equilibrium will be local asymptotically stable whenever the expected net benefits from behaving corruptly in a population that consists only of honest and whistle-blowing agents are lower than the benefits from behaving honestly, i.e., the government's salary. It should be noted that, in this case, the corrupt player confronts two expected costs: the expected costs of being detected by an external enforcing governmental agency and the expected costs of being detected and reported to the government by a whistle-blower from within the organization. From this latter type of cost, typically absent in the traditional economic literature on corruption, makes it clear that the stability of the non-corruption equilibrium crucially depends upon the *initial population share ascribing to the whistle-blowing strategy*, which allows us to infer some additional remarks.

Government organisations with a large initial population of honest bureaucrats that not only do not accept bribes but that are also willing to monitor and report corrupt behaviour ensure higher expected costs for corrupt agents than a baseline case in which whistle-blowing is not a common practice. In this context, even if the probability of being detected by an external enforcing governmental agency and, therefore, the associated expected costs are rather small, the non-corruption equilibrium can still be local asymptotically stable due to the expected costs inflicted on corrupt agents by whistle-blowers.

Our results also indicate the relevance of history and initial conditions when explaining the stability of the corruption or non-corruption equilibria. We conclude that a low initial population of whistle-blowers may imply that the non-corruption equilibrium, in which no corrupt agents are present in the population, may never be stable in the future in a specific organization. Clearly, in order to determine the causes of a low initial population of whistle-blowers, we have to consider the historical background in the analysis. In other words, the current state of affairs must be necessarily analysed considering related events that occurred in the past and somehow determine the present situation. Here, initial conditions are crucial and even small differences may imply widely differing outcomes.

As the empirical evidence on whistle-blowing typically points out, in the real world, whistle-blowing behaviour within organizations is the exception rather than the rule. This is consistent with low populations of whistle-blowers, which may help explain why it is so difficult for public organizations to avoid corrupt agents in the population and how corrupt activities can go on for a long time without being detected or reported by public officials from inside the organization. Consequently, we also study different control instruments that governments could implement to motivate and promote whistle-blowing behaviour in an organization.

In particular, we infer from our model that the local asymptotical stability of the *corruption equilibrium* is contingent on the environment for whistle-blowing behaviour. If the costs associated with whistle-blowing are greater than or equal to the associated benefits, the corruption equilibrium can become very stable. Indeed, if the costs of monitoring a population of only corrupt agents and the transaction costs associated with actually reporting a corrupt government official are high in comparison with the expected benefits of whistle-blowing behaviour, then corruption can be very hard to break. This can be especially so in the total absence of any incentives to whistle-blowers and whenever the probability of being detected by an external enforcing agency is rather small.

By contrast, if the environment for whistle-blowing behaviour is rather favourable, then the corruption equilibrium will only be local asymptotically stable whenever the expected net benefits of behaving corruptly in a population that consists only of corrupt players are greater than the expected net benefits of behaving honestly and performing whistle-blowing activities. This condition for the stability of corruption brings several new elements to the discussion. Additional control instruments for the government to consider apart from the wage rate, the monitoring system, and the penalty include the transaction costs associated with reporting a corrupt public official to the government, monitoring costs, and incentives for whistle-blowers that effectively detect and sanction corrupt agents. Clearly, these alternative control instruments for the government imply new challenges in terms of policy, some of which were addressed in the paper.

If the probability that a whistle-blower will detect a corrupt public official is greater than the probability of detection by an external agency from the government, then the effect of whistle-blowing on the stability of corruption can be quite significant. This is especially true when considering that the more traditional control instruments available to governments for fighting corruption are extremely dependent on the ability of the government to detect corrupt officials. Clearly, an inside member of the organization is more likely to spot corruption, and spot it at an earlier stage, than an external enforcing agency from the government.

All these economic variables can be very important for the government to consider as additional control mechanisms to promote whistle-blowing behaviour. Nonetheless, the potential effect of these instruments on the stability of corruption is not so clear. In particular, we studied the stability of corruption when the economic mechanisms used to promote whistle-blowing behaviour in an organization also encourage *corrupt* players to "blow the whistle". In order to model this effect, we deviated from the basic set-up and assumed the following types of behaviour within an organization: (i) honest behaviour; (ii) corrupt whistle-blowing behaviour; and (iii) honest whistle-blowing behaviour. This analysis showed that there are cases in which the government's control instruments aimed at motivating whistle-blowing behaviour can be counter-productive, making corruption more stable within an organization. Specifically, a high level of economic incentives for whistle-blowers can increasingly induce corrupt agents to become what we have called *corrupt whistle-blowers*, which, in turn, can make honest behaviour and honest whistle-blowing behaviour less economically attractive.

Finally, we also show that, in order to know in what contexts economic incentives for whistle-blowers can contribute more effectively to make corruption less stable within an organization, it is crucial to understand *the nature of monitoring costs*. Thus, in contexts in which marginal monitoring costs are a diminishing function of the corrupt bureaucrat population, economic incentives for whistle-blowers can make corrupt behaviour less economically attractive and, therefore, the corruption equilibrium becomes less stable in the organization. By contrast, whenever marginal monitoring costs are an increasing function of the corrupt bureaucrat population, economic incentives for whistle-blowers can make corrupt behaviour more attractive in economic terms and, therefore, make the corruption equilibrium more stable in the organization. Consequently, a policy recommendation in order to adequately design economic incentives for whistle-blowers would be to investigate in detail monitoring costs in different settings. This is clearly a matter to be resolved empirically.

6. References

Abed, G. T. and Gupta, S., eds., (2002) *Governance, Corruption and Economic Performance*, Washington D.C.: International Monetary Fund.

Acemoglu, D. (1995) Reward Structures and the Allocation of Talent, *European Economic Review*, vol. 39, pp. 17–33.

Aidt, T. (2003) Economic Analysis of Corruption: A Survey, *Economic Journal*, vol. 113: F632-F652.

Alford, C. F., (2001) *Whistleblowers: Broken Lives and Organizational Power*, New York: Cornell University. Press.

Andvig, J. C. and Moene, K.O. (1990) How Corruption may Corrupt, *Journal of Economic Behaviour and Organization*, vol. 13(1), pp. 63–76.

Bardhan, P. (1997) Corruption and Development: A Review of Issues, *Journal of Economic Literature*, vol. 35, pp. 1320–346.

Becker, G. and Stigler, G. (1974) Law Enforcement, Malfeasance and the Compensation of Enforcers, *Journal of Legal Studies*, vol. 3, pp. 1–19.

Bowles, S. (2004) *Microeconomics: Behaviour, Institutions and Evolution*, New Jersey: Princeton University Press

Cadot O. (1987) Corruption as a Gamble, *Journal of Public Economics*, vol. 33, pp. 223–44.

Clark, C. (1997) Whistleblowers, *CQ Researcher* 7: 1057-80.

Davis, M. (1999) Professional Responsibility: Just Following the Rules? *Business and Professional Ethics Journal*, vol. 18, pp. 65-87.

Elliston, Frederick E. (1985). *Whistleblowing Research: Methodological and Moral Issues*, Praeger: New York.

Gintis, H. (2000) *Game Theory Evolving: A Problem-Centered Introduction to Modeling Strategic Interaction*, Princeton, New Jersey, University of Princeton Press.

Glazer, M. P. and P. Glazer (1989) *The Whistleblowers,* Basic Books: New York.

Groeneweg, S. (2001) *Three Whistleblower Protection Models: A Comparative Analysis of Legislation in Australia, the United States and the United Kingdom*, Research Directorate, Public Service Commission of Canada: Ottawa.

Heyes, A. and Kapur, S. (2008) An Economic Model of Whistle-Blower Policy, *Journal of Law, Economics, and Organization,* forthcoming.

Hunt, G. (1997) *Whistleblowing,* in *The Encyclopedia of Applied Ethics*, Academic Press: New York.

Jain A. K. (2001) Corruption: A Review, *Journal of Economic Surveys*, vol. 15(1), pp. 71–121.

Jos, P. and Tompkins, M. (1989) In Praise of Difficult People: A Portrait of the Committed Whistleblower, *Public Administration Review* vol. 49: 552-61.

Jubb, P. (1999) Whistleblowing: A Restrictive Definition and Interpretation, *Journal of Business Ethics*, vol. 21: 77-94.

Lui, F. T. (1986) A Dynamic Model of Corruption Deterrence, *Journal of Public Economics*, vol. 31, pp. 215–36.

Miceli, M. and Near, J. (1992) *Blowing the Whistle: The Organizational and Legal Implications for Companies and Employees*, Lexington Books: New York.

Miethe, T. (1999) *Whistleblowing at Work: Tough Choices in Exposing Fraud, Waste and Abuse on the Job*, Westview Press: Boulder, Colorado.

Mookherjee, B. and Png, I. P. L. (1995) Corruptible Law Enforcers: How Should They be Compensated? *Economic Journal*, vol. 105, pp. 145–59.

Murphy, K. M., Shleifer, A. and Vishny, R. W. (1991) The Allocation of Talent: Implications for Growth, *Quarterly Journal of Economics*, vol. 106, pp. 503–30. Reprinted in Shleifer and Vishny (1998) chapter3.

Murphy, K. M., Shleifer, A. and Vishny, R. W. (1993) Why is Rent Seeking so Costly to Growth? *American Economic Review Paper and Proceedings*, vol. 83, pp. 409–4. Reprinted in Shleifer and Vishny (1998) chapter 4.

Rijckeghem, C. van and Weder, B. S. (2001) Corruption and the Rate of Temptation: Do Low Wages in the Civil Service Cause Corruption? *Journal of Development Economics*, vol. 65, pp. 291–307. Reprinted in Abed and Gupta (2002) chapter 3.

Sah, R. K. (1988) *Persistence and Pervasiveness of Corruption: New Perspectives*, Yale Economic Growth Center Discussion Paper 560.Samuelson (1997),

Samuelson, L. (1997) *Evolutionary Games and Equilibrium Selection*, Cambridge MA: MIT Press.

Shleifer, A. and Vishny, R. W. (1998) *The Grabbing Hand, Government Pathologies and Their Cures*, Cambridge, MA: Harvard University Press.

Søreide, T. (2008) Beaten by Bribery: Why not Blow the Whistle?, *Journal of Institutional and Theoretical Economics / Zeitschrift für die Gesamte Staatswissenschaft* vol. 164 (3): 407-428.

Stephan, A. (2008) Beyond the Cartel Law Handbook: How Corruption, Social Norms and Collectivist Business Cultures Can Undermine Conventional Enforcement Tools. *ESRC Centre for Competition Policy Working Paper* No. 08-29. Available at SSRN: http://ssrn.com/abstract=1277205.

Taylor, P. and Jonker, L. (1978) Evolutionary Stable Strategies and Game Dynamics, *Mathematical Biosciences*, Vol. 40, pp.145-156.

Tirole, J. (1996) A Theory of Collective Reputations (with applications to the Persistence of Corruption and to Firm Quality), *The Review of Economic Studies*, vol. 63, pp. 1–22.

Van Damme, E. (1994) Evolutionary Game Theory, *European Economic Review*, vol. 38: 847-858.

Vega-Redondo, F. (1996) *Evolution, Games and Economic Behaviour*, Oxford: Oxford University Press.

Villena, MG and Villena, MJ (2004) Evolutionary Game Theory and Thorstein Veblen's Evolutionary Economics: Is EGT Veblenian? *Journal of Economic Issues*, 38(3): 585-610.

Weibull, J. (1996) *Evolutionary Game Theory*, London, England: The MIT Press.