University of Innsbruck

Public beliefs and corruption in a repeated psychological game

Loukas Balafoutas

Working Papers in Economics and Statistics

2009-01

eeecon
Research Platform
Empirical and Experimental Economics

University of Innsbruck
http://eeecon.uibk.ac.at/

Public beliefs and corruption in a repeated psychological game *

Loukas Balafoutas [a]

Abstract

This paper investigates the role of guilt aversion for corruption in public administration. Corruption is modeled as the outcome of a game played between a bureaucrat, a lobby, and the public. There is a moral cost of corruption for the bureaucrat, who is averse to letting the public down. We study how the behavior of the lobby and the bureaucrat depend on perceived public beliefs, when these are constant and when they are allowed to vary over time. With time-varying beliefs, corruption is more likely when the horizon of the game is relatively long and when public beliefs are initially low and are updated fast.

November 10th, 2010

[Revised version forthcoming in the *Journal of Economic Behavior and Organization*]

---

[a] University of Innsbruck, Department of Public Finance. Address: Universitätsstrasse 15, A-6020 Innsbruck, Austria. Email address: loukas.balafoutas@uibk.ac.at.

1. Introduction

The economic literature on corruption typically assumes that government officials rationally weigh their expected monetary costs and benefits from corruption, and accordingly decide whether to behave corruptly or not. Moreover, the principal-agent framework is typically used for a structural analysis of corruption (see, for example, Becker and Stigler, 1974; Rose-Ackerman, 1975, 1978; Klitgaard, 1988, 1991; Shleifer and Vishny, 1993; Acemoglu and Verdier, 2000). The resulting implications for policymakers are that corrupt behavior can be curbed by designing appropriate incentive schemes, which increase the costs of corruption for government officials (penalties, effective monitoring), reduce its benefits (limits to the discretionary power of officials), or increase the benefits of honest behavior (efficiency wages, outcome-contingent contracts).[1]

In this paper we focus on a different factor that can affect corrupt behavior. We study the role of emotions for corruption in public administration and suggest that corruption may depend on the beliefs of the public in the context of a repeated psychological game. Psychological games (see Geanakoplos *et al*., 1989; Battigalli and Dufwenberg, 2009) incorporate emotions in the motivation of players by including in their utility not only material payoffs, but also higher-order beliefs. The inclusion of beliefs in utility can lead to emotions such as guilt, embarrassment, disappointment, anger, surprise, or gratitude. In our model we consider guilt aversion as a driving factor for the behavior of bureaucrats, and in order to do this we extend their utility function beyond monetary costs and rewards and include in it the psychological cost of guilt that arises from corrupt behavior. This idea is already mentioned in Klitgaard (1991), who argues that raising the "moral costs" of corruption can be one of five broad sets of tools to control corruption. These "moral costs" of corruption refer mainly to organizational cultures and codes of ethics, something that is applicable to bureaucracies.[2]

Guilt aversion is defined in Charness and Dufwenberg (2006) as the emotion that arises when a player *"believes he hurts others relative to what they believe they will get"*. Further theoretical models on guilt include Dufwenberg (2002), Battigalli and Dufwenberg (2007), while experimental evidence on the effect of guilt on decision-making is abundant (see, for instance, Dufwenberg and Gneezy, 2000; Guerra and Zizzo, 2004; Charness and Dufwenberg, 2006; Dufwenberg *et al*., 2006; Bacharach et al., 2007; Battigalli and

---

[1] On the question of which factors can help reduce the level of corruption, see also Mookherjee and Png (1995), Margit and Shi (1998), Waller *et al*. (2002).

[2] See also van Winden (2007:57) for a discussion on the importance of incorporating emotions in the domain of public choice: *"In my view, by now sufficient evidence exists to replace homo economicus by homo sapiens, a boundedly rational and emotional agent, in the private sector as well as the public sector."*

Dufwenberg, 2007; Reuben *et al*., 2009). Corazzini *et al*. (2007) offer experimental evidence in support of the hypothesis that politicians are motivated by guilt aversion from letting the electorate down. Although their setting is based on electoral competition, the results are important for us because they link the behavior of politicians to their beliefs about the expectations of the electorate, and therefore strengthen the case for a model that adds guilt aversion to the study of corruption in the domain of public administration.

Our stylized model of corruption in public administration shows that, when the bureaucrat is averse to letting the public down, corruption is less likely to emerge in equilibrium compared to a version of the game without psychological payoffs. The model predicts that -under certain assumptions- corruption can lead to more corruption, if the self-fulfilling beliefs of the public enter a downward trajectory. Whether or not a society becomes entrapped in such a corruption spiral depends on a number of factors: corruption is more likely when the psychological component in the bureaucrat's utility is strong relative to the monetary incentives (i.e., when he is very guilt averse), when public beliefs are initially low and are updated fast, and when the expected horizon of the game is relatively long.

In broader terms, the paper adds to the literature that studies the role of beliefs and emotions in decision-making, by considering the impact of an inactive third party on a cooperation game played between two other players, under the assumption of guilt aversion. The model takes a close look at the dynamic properties of beliefs and at how these affect the outcome of the game. Beliefs are initially constrained to be constant over time. We then relax this assumption and move to a version of the model in which players update their beliefs. In such a setting, players must take into account the effect of their actions on others' beliefs in the future; this introduces an additional element of strategic behaviour.

In another application of psychological game theory to the field of public administration and corruption, Huang and Wu (1994) look at a psychological game of trust, where a player's remorse from betraying another player's trust is proportional to their expectations. The central insight of Huang and Wu (1994) is that the expectation-dependent moral cost of betrayal leads to a multiplicity of equilibria, some of which feature (at least some degree of) honest behavior. The authors use corruption in bureaucratic organizations as their central example: bureaucrats may act less corruptly if they believe that the public has high expectations of them. Then, different equilibria with their corresponding sets of beliefs can be viewed as alternative social norms. Compared to Huang and Wu (1994), we endogenize the benefit from corruption by introducing a bribe-paying lobby. Moreover, as the bribe can vary over time, so can the moral cost of corruption. Another paper assuming that

bureaucrats are averse to letting the public down is Varoufakis (2006), who studies the evolution of corruption in bureaucracies with the added twist of linking it to political participation.

2. A model of corruption in public administration

2.1. Players, structure and beliefs

Consider the following three-person game. Player A is a bureaucrat, player B is (some representative of) a business or lobby, and player C is the public.[3] The lobby can bribe the bureaucrat with the purpose of eliciting some favor from him (e.g., a procurement contract with favorable terms), which amounts to corruption. If the bribe is high enough, the lobby may be able to persuade the bureaucrat to act corruptly, which however generates a negative externality for the public. Consistent with the widely used definition of corruption as the *"misuse of public office for private gain"* (Rose-Ackerman, 1999), the bureaucrat can thus choose an action that hurts the public but ensures a higher payoff for him.

To formalize these ideas, let there be two distinct sets of monetary payoffs, called the high and the low set, $\Pi = \{H, L\}$, with $H=(a_H, b_H, c_H)$ and $L=(a_L, b_L, c_L)$, where a, b and c are the payoffs of players A, B and C respectively. Player A determines the outcome of the game by choosing either H or L. Payoffs are such that $a_L<a_H, c_L<c_H, b_L>b_H, a_H-a_L<b_L-b_H$, $a_H+b_H+c_H > a_L+b_L+c_L$. Hence, H gives a relatively high payoff to players A and C, whereas L is more favorable than H to player B but at the same time lowers total welfare.

Before A makes his choice, B has the option of making a transfer to him, which can be understood as a bribe paid to the bureaucrat in order to convince him to choose the payoff set L. Let k be the value of this transfer, $k \geq 0$. The term *corruption* refers to the situation in which players A and B cooperate in order to increase both their payoffs at the expense of player C. Thus, corruption means that B pays a bribe of $k \in [a_H - a_L, b_L - b_H]$ and A chooses the low payoff set L. The condition $a_H-a_L<b_L-b_H$ ensures that there are mutual gains from corruption for A and B. The non-corrupt outcome occurs when A chooses the high payoff set.[4]

Beliefs are defined as follows. Let p denote the probability that A will choose H, and let π denote C's expectation of p. Finally, q is A's expectation of C's expectation of p, in other words his second order belief. In brief, $p=Pr(H)$, $\pi=E^C(p)$, $q=E^A(\pi)$. Player B's expectation of

---

[3] I intentionally refer to a bureaucrat and not a politician. The latter is elected and therefore the public is far from being an inactive player. Bureaucrats, on the other hand, are appointed and their positions are often permanent (although it is true that their careers can depend on political affiliations).

[4] Player A could in theory choose L even if the bribe is lower than (a_H-a_L). This situation is not defined as corruption here, since it reduces A's payoff.

π equals that of player A, i.e., it equals q.[5] Unless otherwise stated, the term "beliefs" will refer throughout the paper to the bureaucrat's second order belief q.

The utility functions of players B and C comprise simply their monetary payoffs. As explained in the introduction, the central assumption of the model is that player A is guilt averse, i.e., he is averse to letting player C down. This is modeled as follows. If A chooses L, his utility includes the monetary payoff and the bribe (if any), but also a psychological cost from betraying C's expectations (see Figure 1). In other words, there is an endogenous benefit to corruption (equal to the bribe), and a dual cost to it. The bureaucrat may have to exert some effort in order to grant the favor or in order to preserve secrecy and avoid getting caught -this cost corresponds to the difference a_H-a_L. Alternatively, this difference can be thought of as the expected cost from being caught (probability of detection multiplied by the penalty). In addition to this direct cost, the bureaucrat suffers a moral cost of corruption. This is equal to γq, where the parameter γ measures the intensity of guilt aversion. The value of γ is common knowledge.[6] The model conditions guilt only on the beliefs of the public, and not on those of the lobby. Section 4 relaxes this assumption and looks at the implications of allowing the bureaucrat to feel guilt when he lets any of the two players down.

Figure 1: Player A's utility
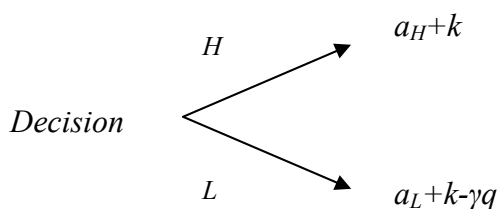
Under our assumption of guilt aversion, the bureaucrat is not a purely selfish money-maximizer and his motivation includes the fact that he feels he must live up to the expectations of the public. This sense of duty and responsibility is stronger, the higher the public's expectations. Consider a bureaucrat who acts corruptly. The term $(-\gamma q)$ in his utility function implies that this bureaucrat will feel less guilt in a society where corruption is widely regarded as very likely than in a society that holds high expectations of him.

---

[5] The assumption that A and B have the same second order beliefs can be relaxed without leading to any qualitative changes in the results, but increasing the complexity of the analysis.
[6] On this assumption, see the discussion in section 3.1.

2.2. The one-shot solution

Given the timing of events, backwards induction reveals that no bribe is paid in equilibrium, since player B correctly anticipates that A's dominant strategy is to choose *H* for every possible amount of the bribe. This is a standard commitment problem and corruption cannot be sustained in equilibrium. In this example, the psychological assumption regarding the bureaucrat's utility does not affect the outcome. Indeed, guilt aversion pushes him even stronger towards choosing *H*, since the choice of *L* generates an additional psychological cost. However, beliefs can make a difference when the game is indefinitely (even if finitely) repeated.

2.3. The repeated game

Let the above game be played repeatedly and suppose that after every repetition the probability that the game will be played once more equals $r \in (0,1)$, where *r* is constant across periods. In this dynamic context, players may be able to establish corruption by adopting some conditional collusive strategy. In particular, consider the possibility that B adopts the following grim trigger strategy (called S_B): *"Pay a positive bribe x in the first period, and then only keep paying the same bribe provided A has colluded in the previous period; switch permanently to a zero bribe if in any period A does not collude"*. Turning to the bureaucrat, he may also adopt a trigger strategy of the following sort: *"Collude as long as B has paid at least y, never collude if B pays less than y in any period"*.

Given these trigger strategies, an infinite number of equilibria are possible, some of which support corruption (those in which $x \geq y$). The problem is that there is no natural way of selecting among these equilibria.[7] Although there are infinite intertemporal strategies that can be adopted by the players and supported in equilibrium, we shall confine ourselves in what follows to certain focal strategies. In particular, the analysis will focus here on the case where B plays S_B and A plays the following simple strategy, called S_A: *"Collude as long as the payoff from corruption is at least equal to the payoff from no corruption"*. These strategies correspond to a situation where the lobby is able to extract the maximum benefit from corruption, while keeping the bureaucrat indifferent between corruption and no corruption. Moreover, when public beliefs are updated, these strategies have the interesting feature that they generate time-varying bribes, which are functions of those beliefs. In the appendix we discuss another extreme case, where strategies are such that the bureaucrat is

---

[7] As Binmore and Shaked (2010:88) point out, *"the case of multiple Nash equilibria is hard, because most game theorists regard the equilibrium selection problem as unsolved"*.

able to extract the maximum possible bribe from the lobby. The choice of these two cases will allow us to pin down the respective corruption equilibria with extreme allocations: one where the lobby is able to capture the entire surplus from corruption, and one where the bureaucrat captures the entire surplus. For the sake of simplicity we will only study these two cases, although any intermediate sharing of the surplus is possible in some equilibrium of the repeated game.

Let q, the bureaucrat's second-order belief, be constant throughout the horizon of the game. We also require that q be correct in equilibrium, i.e., that the second-order belief correspond to reality. This assumption is necessary in order to use the concept of a psychological Nash equilibrium (henceforth PNE, see Geanakoplos *et al*., 1989) for the solution of the game. Both of these assumptions will be relaxed in the next section.

Player B plays S_B and player A plays S_A. As we have seen, this strategy amounts to A colluding provided that the bribe covers at least the difference between a_H and a_L, and it is the strategy that gives the lowest collusive payoff to A and the highest collusive payoff to B. Given that B plays S_B each period, and given that none of the parameters change over time, she will pay the same bribe k every period if there is corruption. Accordingly, A will either collude in every period or he will never collude. The associated respective total expected payoffs for A are: $(a_L + k - \gamma q)/(1-r)$, $k + a_H/(1-r)$. If A honours B's trust and colludes, then each period his utility equals the low monetary payoff a_L plus the bribe minus the cost from guilt. If he chooses not to collude, then his utility is high in the first period, but in all subsequent periods he receives only a_H. Comparing the two payoffs, it follows that player A will collude as long as: $r \geq [(a_H - a_L) + \gamma q]/k$. *(1)*

This expression conditions the sustainability of corruption on the effect of guilt (γq), the monetary loss that player A suffers if he chooses *L (a_H-a_L)*, the expected duration of the interaction and the amount of the bribe. For A to collude, the bribe and the expected horizon of the game must be sufficiently high with respect to the difference in monetary payoffs and to the psychological cost of disappointing the public.

Lemma 1

Given S_A and S_B, if corruption is the equilibrium of the game B pays the following bribe k^* every period: $k* = (a_H - a_L + \gamma q)/r$

This follows from *(1)*.

A bribe of k^* compensates A for the utility loss caused by the lower monetary payoff and the cost of guilt. A high continuation probability r reduces the amount of the bribe. This accounts for the fact that the bribe is paid in advance, so that in any given period T the benefit to A from corruption lies in the future while the cost is borne starting at T.

Proposition 1

Given that A and B follow strategies S_A and S_B respectively, the game has two possible equilibria:

(i) The corruption equilibrium, with $\Pi=L$ in every period, $p=q=0$, and $k^*=(a_H-a_L)/r$.

(ii) The no corruption equilibrium, with $\Pi=H$ in every period, $p=q=1$, and $k^*=0$.

If $(a_H - a_L + \gamma)/r \leq b_L - b_H$, the equilibrium is necessarily (i)

If $(a_H - a_L)/r \geq b_L - b_H$, the equilibrium is necessarily (ii)

Otherwise, both equilibria are possible.

Proof: $(a_H - a_L + \gamma)/r \leq b_L - b_H$ means that, from *(1)*, B can always pay a bribe high enough (and still lower than b_L-b_H) to sustain corruption, even if perceived expectations are at the highest possible level *(q=1)*. Knowing this, the public is convinced that A and B will collude, q is actually zero and A confirms these expectations by choosing *L*. On the other hand, $(a_H - a_L)/r \geq b_L - b_H$ means that, from *(1)*, B cannot pay a bribe lower than b_L-b_H and still high enough to sustain corruption, even if $q=0$. Therefore, everyone expects the no-corruption equilibrium *(p=q=1)* and expectations are again confirmed.

The above equilibria satisfy the definition of a PNE: players' strategies are optimal and beliefs correspond to actual play. There is a multiplicity (at least within certain parameter values), meaning that we could end up in any of the two equilibria.[8] In this repeated game, the psychological element in the bureaucrat's motivation makes corruption less likely to emerge in equilibrium. To see this, note that in the absence of guilt aversion (i.e., when $\gamma=0$) the necessary and sufficient condition for corruption is: $(a_H - a_L)/r \leq b_L - b_H$ *(2)*

This is the solution to the standard, non-psychological version of the game. Comparing *(2)* with the conditions given in Proposition 1 reveals that, for certain parameter values, corruption is always the outcome when $\gamma=0$ but not necessarily when $\gamma>0$.

---

[8] Of course, as in any other infinitely repeated game of interest, many other equilibria exist if we consider strategies other than S_A and S_B, inclusive equilibria in mixed strategies (which necessarily involve less cooperation than the pure strategy equilibrium characterized in Proposition 1).

3. Time-varying beliefs

3.1. General

So far it has been assumed that beliefs are constant and that they correspond to the equilibrium outcome, which is reached in the first period and does not change over the horizon of the game. These assumptions are now relaxed and beliefs are allowed to vary over time. Before we look at the specifics of the model, we begin with a remark regarding equilibrium. The model in this section will posit a specific behavioral assumption regarding the evolution of beliefs (in the form of an updating process), generating beliefs that are not necessarily correct in equilibrium. This means that the model no longer satisfies the assumptions required for the application of PNE or sequential equilibrium (SE) as solution concepts.[9] In light of this fact, we will consider a state in which players A and B are in equilibrium in their corruption game, but the inactive player C may hold out-of-equilibrium beliefs about its outcome.

One way to motivate the above ideas is to consider the possibility of asymmetric information in the model. It was assumed that all the parameters (payoffs and psychological factors) are common knowledge; this assumption appears unrealistic in many settings, particularly in the context of psychological games.[10] If player C knows all the parameters of the model as well as the strategies that A and B follow, beliefs will jump to a corner and we will end up with one of the equilibria described in Proposition 1. If, instead, information is incomplete, it seems appropriate to not constrain beliefs to be correct in equilibrium, but to allow them to take interior values. For this, it is sufficient to think that C may not know some of the parameters of the model (for instance the psychological parameter γ), or that he simply does not know the particular strategies S_A and S_B that the other players follow, which would mean that from C's point of view anything could be an equilibrium of the repeated game.

On a related issue, the analysis of the game hinges crucially on the assumption that B knows how guilt averse A is, and thus, she can pay the appropriate bribe. This is admittedly a strong assumption- even though it is common in the literature on psychological games. To motivate it, one can think that the lobby invests some resources in order to find out the extent to which a given bureaucrat is prone to corruption (e.g., based on his behavior in other

---

[9] Both the SE (Battigalli and Dufwenberg, 2009) and PNE require that beliefs of any order be correct in equilibrium. See Battigali and Dufwenberg (2009: 16), for their consistency assumption as well as its analogy to PNE.

[10] *"Unless one models interaction within a family or amongst friends, it is probably not realistic to assume that players know each other's psychological propensities"* (Battigalli and Dufwenberg 2009:27). Also, Rabin (1993) motivates the study of psychological games with incomplete information, where a player's knowledge of other players' motives may be limited.

positions in the past). The public is not organized and able to undertake collective action in a similar way as the lobby, and therefore may not know the true value of γ.

3.2. The model with updated beliefs

Beliefs are initially at some exogenous level q_0 and they can change over time $(q=q_t)$: the public updates its belief about A's choice each period taking into account A's choice in the preceding period, and players A and B update their perceived beliefs in exactly the same way.[11] Before imposing a specific process, we begin with a general case where the only assumption is that beliefs fall every time player A chooses L, and that they rise every time he chooses H:

$$\begin{cases} q_{t+1} > q_t, & \text{if } \Pi_t = H \text{ and } q_t < 1 \\ q_{t+1} < q_t, & \text{if } \Pi_t = L \text{ and } q_t > 0 \end{cases} \tag{3}$$

Following the discussion in section 3.1, we define an equilibrium of our corruption game with updated beliefs as a pair of strategies S_A and S_B of players A and B, and a belief q_t of player C, such that: (i) in every period, S_A and S_B are best responses to each other, and (ii) q_t is updated according to (3).

Proceeding as in section 2, we can determine the amount of the bribe that in every period makes player A indifferent between always colluding and never colluding (from that period onwards). This is given by: $k^*_t = (a_H - a_L + \gamma q_t)/r$ \hfill (4)

Proposition 2

Given (3), S_A and S_B, the unique optimal sequence of bribe payments made by player B in the corruption equilibrium is k^*_t as given by (4).

Proof: In any period T throughout the game, no bribe $k_T > k^*_T$ will be paid; since k^*_T is high enough to make player A prefer the collusive outcome subject to k^*_t being offered in all subsequent periods, a higher bribe unnecessarily reduces B's payoff. On the other hand, no $k_T < k^*_T$ is offered in period T either. Such a bribe is lower than what is needed to compensate player A, so the latter will not collude in period T with $k_T < k^*_T$, unless he expects to be compensated for the loss in T in some future period(s). But such an expectation is not rational and any such commitment by player B would not be credible, since we have just seen that k_t will never exceed k^*_t. Thus, k_T must equal exactly k^*_T.

---

[11] The assumption that A and B have correct second order beliefs is not crucial. All that is required is that A and B have the same second order beliefs. See also footnote 5 on this.

Given (3), S_A and S_B, the sequence of optimal bribe payments in the corruption equilibrium is strictly decreasing over time: $k_t^* > k_{t+1}^*$

This follows immediately from *(3)* and *(4)*.

This corollary is saying that, in this particular setting, the bribe that the lobby must pay falls over time. Hence, corruption becomes cheaper for the lobby each period. The reason why this happens is that the bribe must compensate the bureaucrat for the cost of guilt, and this cost falls each period as expectations become lower. This result implies that, by lowering public expectations and eroding confidence in public administration over time, corruption paves the way for more corruption.[12]

We will now examine the conditions under which the corruption equilibrium emerges. In any period, paying a bribe in exchange for the choice of *L* is profitable for B if $k_t < b_L - b_H$, i.e., if her cost (bribe) is lower than her benefit. Moreover, from Corollary 1, if this action is profitable in the first period, it will be profitable in all subsequent periods. The intuition is that expectations fall over time and the required compensation diminishes as guilt becomes weaker. This leads to the fact that an equilibrium with corruption in every period is definitely profitable for player B if: $(a_H - a_L + \gamma q_0)/r \leq b_L - b_H$ *(5)*

This condition ties the emergence of corruption to initial beliefs. But that is not the whole story. Condition *(5)* is sufficient for corruption to be profitable for B, but it is not necessary. It may be the case that B wants to sustain corruption even if it is unprofitable for her in the first period. The reason is that, as beliefs fall, k_t^* will also fall over time so that the total expected return from always paying the bribe may be positive. In other words, player B might make a monetary sacrifice today in the expectation that she will be compensated for it tomorrow, since the effect of this sacrifice will be to drive expectations down. In order to be able to say something about the outcome of the game in that case, we must assume a specific dynamic process for beliefs and assess the benefits and costs of setting $k_t = k_t^*$ as in *(4)* throughout the horizon of the game. Let beliefs be updated according to the following process:

---

[12] Another paper in which corruption leads to more corruption is Tirole (1996). In a principal-agent model, Tirole (1996) shows how an economy can become "locked" in a high corruption steady state when the principal's source of information about the agent is the agent's group reputation.

$$\begin{cases} q_{t+1} = q_t + \rho(1-q_t), & if \ \Pi_t = H \ and \ q_t < 1 \\ q_{t+1} = q_t - \rho q_t, & if \ \Pi_t = L \ and \ q_t > 0 \end{cases} \qquad (6)$$

This process specifies that in any given period beliefs will move either upwards, if there was no corruption in the previous period, or downwards if there was corruption.[13] The parameter $\rho \in [0,1]$ measures the speed of beliefs adjustment.[14] If $\rho = 1$, beliefs jump to a corner after the first period. In the opposite case where $\rho=0$, beliefs do not move over time and remain at their initial value of q_0 (as in the previous section). This process captures the main features of the adjustment of beliefs that we have in mind, although one can think of alternative specifications.

Solving the difference equations in *(6)* gives the following expressions for beliefs at any point in time as a function of initial beliefs:

$$\begin{cases} q_t = q_0(1-\rho)^t + 1 - (1-\rho)^t, & in \ the \ no \ corruption \ equilibrium \\ q_t = q_0(1-\rho)^t, & in \ the \ corruption \ equilibrium \end{cases} \qquad (7)$$

From *(4)* and *(7)*, we obtain the level of the bribe in any period t:

$$k_t^* = [a_H - a_L + \gamma q_0(1-\rho)^t]/r \qquad (8)$$

Proposition 3

When beliefs follow the dynamic process given in (6), and given S_A and S_B, the game has two possible equilibria:

(i) If $\dfrac{(a_H - a_L)}{r} + \dfrac{\gamma q_0(1-r)}{r[1-r(1-\rho)]} \le b_L - b_H$, $\qquad (9)$

the outcome is corruption in every period with $k_t = k^*_t$ as given in (8).
(ii) If (9) does not hold, corruption never occurs.

Proof: The corruption equilibrium emerges as long as the expected total value of the bribe payments throughout the game does not exceed the expected value of the total benefit of corruption to player B:

$\sum_{t=0}^{\infty} k_t^* r^t \le \sum_{t=0}^{\infty} (b_L - b_H) r^t$, which translates to *(9)*. In such a case, player B will always pay k^*_t, because

to her this is the least costly way of sustaining corruption. On the contrary, if *(9)* does not hold, player B never pays a bribe and player A always chooses *H*.

---

[13] The first part of the process that refers to the case where there is no corruption does not enter into any of the calculations and does not affect the results, because in the no corruption equilibrium the level of beliefs does not matter.
[14] In fact, the speeds of upward and downward adjustment should not be constrained to be equal. This would account, for example, for the possibility that confidence might take a longer time to build than to destroy. Here we use only one parameter ρ for simplicity, since the upward path of beliefs is irrelevant to the outcome of the game.

3.3. Discussion: Factors that facilitate corruption

Starting from an initial level of beliefs and observing the incidence -or not- of corruption, the public updates its expectations every period. Proposition 3 predicts that, as long as the exogenous parameters do not change, corruption will occur either in all periods or in none. It therefore also predicts that trust in bureaucrats will keep moving, according to *(6)*, either towards very high levels or towards complete disillusionment where everyone expects them to act corruptly. If that is the case, lobbies will eventually be able to capture bureaucrats with only small amounts of bribes. The intuition for this is that a bureaucrat who knows that the public considers him corrupt suffers a low moral cost when he actually behaves in a corrupt manner, so that he is eventually caught in a circle of self-fulfilling expectations.

Compared to the version of the game with constant beliefs, there is now no multiplicity of equilibria and the outcome is determined by *(9)*. This condition reveals that corruption is more likely to emerge in equilibrium when: *(i)* the stakes of corruption (b_L-b_H) are high and the cost to the bureaucrat (a_H-a_L) is low, *(ii)* guilt aversion is relatively weak (low γ), *(iii)* initial beliefs q_0 are low, *(iv)* the expected horizon of the game is long: as r increases, the effect of guilt diminishes relative to the other factors that determine the outcome of the game,[15] making corruption less costly for the bureaucrat. In policy terms, this means that the frequent rotation of public servants can be an effective instrument against corruption, *(v)* Beliefs are updated fast (high ρ): if the public is quick to become disillusioned by a bureaucrat, then the latter is more likely to act corruptly in the first place. On the contrary, a bureaucrat is less likely to disappoint the public if beliefs are relatively resilient. The intuition behind this result is that, if expectations decline fast following the incidence of corruption, then the psychological cost of guilt also diminishes fast and the bribe that is required to establish corruption falls. Note, also, that the model predicts that lobbies are likely to act strategically and offer very high bribes at initial stages of the corruption relationship, with the purpose of marring the image of public officials and making corruption easier and cheaper in the future. This strategy is more profitable, the higher is the speed of dynamic adjustment of beliefs.

The last two points *(iv)* and *(v)* highlight the fact that allowing beliefs to vary over time leads to a *time-varying cost of guilt*. If the public's expectations fall during the game, then so does the psychological cost to the bureaucrat of betraying these expectations, and vice versa. Summing up the main idea of the model, high perceived expectations make corruption

---

[15] This is because q_t falls over time, whereas the other parameters remain constant.

harder to sustain because they imply a high guilt-related disutility for the bureaucrat; on the contrary, low expectations can actually *lead* to corruption. This highlights the self-fulfilling property of public beliefs in this game.

4. Guilt from letting both players down

A question that may arise concerning the specification of the game is the following: why does the bureaucrat feel guilt *only* when he lets the public down? In other words, why does he play a psychological game with one of the players and a standard (neoclassical) game with the other? To motivate this assumption, one may evoke social and psychological arguments. It is natural to assume that a bureaucrat has some sense of mission and responsibility towards the public that he is supposed to serve, but not towards a lobby. Nevertheless, this section modifies the model of section 3.2 by adding guilt from letting the lobby (player B) down. In particular, let us assume that, if a positive bribe is paid and player A selects H, he suffers guilt. In this context, guilt does not arise from the choice of H per se, but from the fact that A fails to respond to B's trust.[16]

The only difference that this additional psychological assumption makes in the analysis of the game is that, if in period t player B pays a bribe and player A selects H, the latter's payoff becomes $U^H_t = a_H + k - \xi\mu_t$, where ξ measures the strength of guilt aversion and μ_t is A's second-order belief, defined in the same way as q_t: it is A's expectation of B's expectation of the probability that A will select L. Note however that, given S_B, guilt can only occur once, in the first period when A does not collude. This is because in any subsequent period B will switch to a zero bribe strategy. Thus, the problem facing the bureaucrat is now not the same in every period. If he lets player B down he will have a lower utility in that particular period but will afterwards return to the high monetary payoff a_H. For simplicity, let $\mu_t = \mu$: since beliefs do not matter anymore after A selects H for the first time, there is no need for an updating rule similar to *(3)* or *(6)*. Given S_B and S_A, the optimal sequence of bribes in the corruption equilibrium is now: $k^*_t = [a_H - a_L + \gamma q_t - \xi\mu(1-r)]/r$ *(10)*

The difference between k^*_t in *(4)* and in *(10)* is the additional term $-\xi\mu(1-r)$, which reduces the level of the bribe that is required to sustain corruption. Compared to the game of section 3.2, it is now relatively cheaper for the lobby to influence the bureaucrat, the more so the shorter is the horizon of the game. In fact, with the effect of guilt now going in two different directions, it is the relative disutility of letting the two players down that matters for the outcome of the game. The term *(1-r)* reflects the fact that guilt from letting B down can

---

[16] If player A suffers guilt whenever he chooses H irrespective of whether a bribe has been paid, the results are almost identical to the ones that are obtained here.

only occur once, so that only part of it enters A's utility seen from any point t in time. Then, given *(6)*, the condition for corruption to be the equilibrium outcome becomes:

$$\frac{(a_H - a_L)}{r} + \frac{\gamma q_0 (1 - r)}{r[1 - (1 - \rho)r]} - \frac{\xi \mu (1 - r)}{r} \leq b_L - b_H \qquad (11)$$

In addition to the findings of section 3, we see here that corruption is more likely when player B's perceived beliefs *(μ)* are high and when A's guilt from disappointing B is strong (high *ξ*). The effect of the continuation probability *r* on the incidence of corruption is no longer clear-cut, because it works in two opposite directions through the second and third term on the left-hand side of *(11)*.

5. Conclusion

The economic literature on corruption has largely neglected the role of emotions as a driving factor for the behavior of actors involved in corrupt activities. In this paper we consider how guilt aversion affects the behavior of a bureaucrat, and we do this by modelling corruption in public administration as the outcome of a game of collusion between a bureaucrat and a lobby, which generates an externality for the public. We have used a stylized model to show how guilt aversion can help mitigate the problem of corruption and how it affects the strategic behavior of lobbies and the likelihood that they are able to sustain corruption by bribing public officials. Some insights regarding the interaction between a bureaucrat, a lobby and the public are that declining public expectations allow the lobby to capture the bureaucrat with lower bribes; that this paves the way for more and more corruption as public beliefs enter a self-fulfilling downward trajectory; finally, that corruption is more likely when public beliefs are updated fast and when the interaction between the bureaucrat and the lobby has a long expected horizon.

The analysis of these issues has relied on specific assumptions regarding the psychological motivation of the bureaucrat, the updating of public beliefs and the common knowledge of certain parameters of the model, and it has concentrated on a focal subset of the strategies that players may adopt in the repeated game. In that sense, the paper leaves open many possibilities for future research, which could lead to a more general framework for the study of this topic.

Appendix: Considering a different set of strategies.

Consider once more the model of section 3.2. Player A may be able to maximize his payoff from corruption through the following trigger strategy, called S'_A: *"Collude as long as the bribe is at least equal to $y^*=(b_L-b_H)$"*. This more aggressive strategy maximizes player A's payoff, provided that player B follows the following simple strategy S'_B, which is similar to S_A: *"Pay a bribe high enough to sustain corruption, as long as the payoff from corruption is at least equal to the payoff from no corruption"*.

The value of y^* is such that player B is indifferent between corruption and no corruption. Given S'_A, S'_B, and *(3)*, the equilibrium of the game with updated beliefs is the following: B pays each period a bribe of $k=y^*$ and corruption is the equilibrium. B's total expected payoff from corruption is at its possible lowest level and equal to $b_H/(1-r)$, while A's expected payoff is equal to $(a_L+b_L-b_H-\gamma q_t)/(1-r)$.

In this case we do not have a time-varying bribe as in the solution of section 3.2. The reason is that the bribe is at a maximum and does not fall over time given strategies S'_A and S'_B. The fall in public beliefs over time leads to an increasing payoff for the bureaucrat given that the bribe remains constant, while in section 3.2 it was the lobby that benefitted from the fall in beliefs.

If we consider a wider set of possible strategies by the two players, Corollary 1 on the decreasing bribe payments over time will hold as long as the moral cost of corruption for the bureaucrat, given by γq_t, makes its way into the expression for y (the minimum bribe required by the bureaucrat according to his strategy). Of course, given any such set of strategies, corruption may or may not be the equilibrium outcome, depending on a number of factors captured by an inequality such as *(9)*.

References

Acemoglu D., Verdier T., 2000. The choice between market failures and corruption. *American Economic Review* 90: 194-211

Battigalli P., Dufwenberg M., 2009. Dynamic psychological games. *Journal of Economic Theory* 114: 1-35

Battigalli P., Dufwenberg M., 2007. Guilt in games. *Americal Economic Review* 97: 170-176

Bacharach M., Guerra G., Zizzo D., 2007. The self-fulfilling property of trust: An experimental study. *Theory and Decision* 63: 349-388

Becker G., Stigler G., 1974. Law enforcement, malfeasance, and the compensation of enforcers. *Journal of Legal Studies* 3: 1-19

Binmore, K., Shaked, A., 2010. Experimental economics: Where next? *Journal of Economic Behavior and Organization* 73, 87-100

Charness G., Dufwenberg M., 2006. Promises and partnership. *Econometrica* 74: 1579-1601

Corazzini L., Kube S., Maréchal M., 2007. Towards behavioral public choice: Guilt aversion and accountability in the lab. ISLA Working Paper no. 27

Dufwenberg M., 2002. Marital investments, time consistency and emotions. *Journal of Economic Behavior and Organization* 48: 57-69

Dufwenberg M., Gächter S., Hennig-Schmidt H., 2006. The framing of games and the psychology of strategic choice. CeDEx Discussion Paper no. 2006-20

Dufwenberg M., Gneezy U., 2000. Measuring beliefs in an experimental lost wallet game. *Games and Economic Behavior* 30: 163-182

Geanakoplos J., Pearce D., Stacchetti E., 1989. Psychological games and sequential rationality. *Games and Economic Behavior* 1: 60-79

Guerra, G., Zizzo, D., 2004. Trust responsiveness and beliefs. *Journal of Economic Behavior and Organization* 55: 25-30.

Huang P., Wu, H., 1994. More order without more law: A theory of social norms and organizational cultures. *Journal of Law, Economics and Organization* 10: 390-406

Klitgaard R., 1988. *Controlling Corruption*, University of California Press, Berkeley

Klitgaard R., 1991. Gifts and bribes. In: Richard Zeckhauser (ed.), *Strategy and Choice*, MIT Press

Marjit S., Shi H., 1998. On controlling crime with corrupt officials. *Journal of Economic Behavior and Organization* 34: 163-172

Mookherjee D., Png I., 1995. Corruptible law enforcers: How should they be compensated? *Economic Journal* 105: 145-159

Reuben E., Sapienza P., Zingales L., 2009. Is mistrust self-fulfilling? *Economics Letters* 104: 89-91

Rabin, M., 1993. Incorporating fairness into game theory and economics. *American Economic Review* 83: 1281-1302

Rose-Ackerman S., 1975. The economics of corruption. *Journal of Public Economics* 4: 187-203

Rose-Ackerman S., 1978. *Corruption: A Study of Political Economy*, New York: Academic Press

Rose-Ackerman S., 1999. Corruption and government: Causes, consequences, and reform. Cambridge University Press

Shleifer A., Vishny R., 1993. Corruption. *The Quarterly Journal of Economics* 108: 599-617

Tirole J., 1996. A theory of collective reputations (with applications to the persistence of corruption and to firm quality). *Review of Economic Studies* 63: 1-22

van Winden F., 2007. Affective public choice. In: J. C. Pardo and P. Schwartz (eds.), *Public Choice and the Challenges of Democracy*, Edward Elgar

Varoufakis Y., 2006. The bonds that impede: A model of the joint evolution of corruption and apathy. *Indian Economic Journal* 54: 84-103

Waller C., Verdier T., Gardner R., 2002. Corruption: Top down or bottom up? *Economic Inquiry* 40: 688-703

University of Innsbruck - Working Papers in Economics and Statistics
Recent Papers can be accessed on the following webpage:

2011-10 **Stefan Borsky, Paul A. Raschky:** A spatial econometric analysis of compliance with an international environmental agreement on open access resources

2011-09 **Edgar C. Merkle, Achim Zeileis:** Generalized measurement invariance tests with application to factor analysis

2011-08 **Michael Kirchler, Jürgen Huber, Thomas Stöckl:** Thar she bursts - reducing confusion reduces bubbles *modified version forthcoming in* American Economic Review

2011-07 **Ernst Fehr, Daniela Rützler, Matthias Sutter:** The development of egalitarianism, altruism, spite and parochialism in childhood and adolescence

2011-06 **Octavio Fernández-Amador, Martin Gächter, Martin Larch, Georg Peter:** Monetary policy and its impact on stock market liquidity: Evidence from the euro zone

2011-05 **Martin Gächter, Peter Schwazer, Engelbert Theurl:** Entry and exit of physicians in a two-tiered public/private health care system

2011-04 **Loukas Balafoutas, Rudolf Kerschbamer, Matthias Sutter:** Distributional preferences and competitive behavior *forthcoming in* Journal of Economic Behavior and Organization

2011-03 **Francesco Feri, Alessandro Innocenti, Paolo Pin:** Psychological pressure in competitive environments: Evidence from a randomized natural experiment: Comment

2011-02 **Christian Kleiber, Achim Zeileis:** Reproducible Econometric Simulations

2011-01 **Carolin Strobl, Julia Kopf, Achim Zeileis:** A new method for detecting differential item functioning in the Rasch model

2010-29 **Matthias Sutter, Martin G. Kocher, Daniela Rützler and Stefan T. Trautmann:** Impatience and uncertainty: Experimental decisions predict adolescents' field behavior

2010-28 **Peter Martinsson, Katarina Nordblom, Daniela Rützler and Matthias Sutter:** Social preferences during childhood and the role of gender and age - An experiment in Austria and Sweden *Revised version forthcoming in* Economics Letters

2010-27 **Francesco Feri and Anita Gantner:** Baragining or searching for a better price? - An experimental study. *Revised version accepted for publication in* _Games and Economic Behavior_

2010-26 **Loukas Balafoutas, Martin G. Kocher, Louis Putterman and Matthias Sutter:** Equality, equity and incentives: An experiment

2010-25 **Jesús Crespo-Cuaresma and Octavio Fernández Amador:** Business cycle convergence in EMU: A second look at the second moment

2010-24 **Lorenz Goette, David Huffman, Stephan Meier and Matthias Sutter:** Group membership, competition and altruistic versus antisocial punishment: Evidence from randomly assigned army groups

2010-23 **Martin Gächter and Engelbert Theurl:** Health status convergence at the local level: Empirical evidence from Austria *(revised Version March 2011)*

2010-22 **Jesús Crespo-Cuaresma and Octavio Fernández Amador:** Buiness cycle convergence in the EMU: A first look at the second moment

2010-21 **Octavio Fernández-Amador, Josef Baumgartner and Jesús Crespo-Cuaresma:** Milking the prices: The role of asymmetries in the price transmission mechanism for milk products in Austria

2010-20 **Fredrik Carlsson, Haoran He, Peter Martinsson, Ping Qin and Matthias Sutter:** Household decision making in rural China: Using experiments to estimate the influences of spouses

2010-19 **Wolfgang Brunauer, Stefan Lang and Nikolaus Umlauf:** Modeling house prices using multilevel structured additive regression

2010-18 **Martin Gächter and Engelbert Theurl:** Socioeconomic environment and mortality: A two-level decomposition by sex and cause of death

2010-17 **Boris Maciejovsky, Matthias Sutter, David V. Budescu and Patrick Bernau:** Teams make you smarter: Learning and knowledge transfer in auctions and markets by teams and individuals

2010-16 **Martin Gächter, Peter Schwazer and Engelbert Theurl:** Stronger sex but earlier death: A multi-level socioeconomic analysis of gender differences in mortality in Austria

2010-15 **Simon Czermak, Francesco Feri, Daniela Rützler and Matthias Sutter:** Strategic sophistication of adolescents - Evidence from experimental normal-form games

2010-14 **Matthias Sutter and Daniela Rützler:** Gender differences in competition emerge early in live

2010-13 **Matthias Sutter, Francesco Feri, Martin G. Kocher, Peter Martinsson, Katarina Nordblom and Daniela Rützler:** Social preferences in childhood and adolescence - A large-scale experiment

2010-12 **Loukas Balafoutas and Matthias Sutter:** Gender, competition and the efficiency of policy interventions

2010-11 **Alexander Strasak, Nikolaus Umlauf, Ruth Pfeifer and Stefan Lang:** Comparing penalized splines and fractional polynomials for flexible modeling of the effects of continuous predictor variables

2010-10 **Wolfgang A. Brunauer, Sebastian Keiler and Stefan Lang:** Trading strategies and trading profits in experimental asset markets with cumulative information

2010-09 **Thomas Stöckl and Michael Kirchler:** Trading strategies and trading profits in experimental asset markets with cumulative information

2010-08 **Martin G. Kocher, Marc V. Lenz and Matthias Sutter:** Psychological pressure in competitive environments: Evidence from a randomized natural experiment: Comment

2010-07 **Michael Hanke and Michael Kirchler:** Football Championships and Jersey sponsors' stock prices: An empirical investigation

2010-06 **Adrian Beck, Rudolf Kerschbamer, Jianying Qiu and Matthias Sutter:** Guilt from promise-breaking and trust in markets for expert services - Theory and experiment

2010-05 **Martin Gächter, David A. Savage and Benno Torgler:** Retaining the thin blue line: What shapes workers' intentions not to quit the current work environment

2010-04 **Martin Gächter, David A. Savage and Benno Torgler:** The relationship between stress, strain and social capital

2010-03 **Paul A. Raschky, Reimund Schwarze, Manijeh Schwindt and Ferdinand Zahn:** Uncertainty of governmental relief and the crowding out of insurance

2010-02 **Matthias Sutter, Simon Czermak and Francesco Feri:** Strategic sophistication of individuals and teams in experimental normal-form games

2010-01 **Stefan Lang and Nikolaus Umlauf:** Applications of multilevel structured additive regression models to insurance data

University of Innsbruck

Working Papers in Economics and Statistics

2009-01

Loukas Balafoutas

Public beliefs and corruption in a repeated psychological game

Abstract
This paper investigates the role of guilt aversion for corruption in public administration. Corruption is modeled as the outcome of a game played between a bureaucrat, a lobby, and the public. There is a moral cost of corruption for the bureaucrat, who is averse to letting the public down. We study how the behavior of the lobby and the bureaucrat depend on perceived public beliefs, when these are constant and when they are allowed to vary over time. With time-varying beliefs, corruption is more likely when the horizon of the game is relatively long and when public beliefs are initially low and are updated fast.