On the cultural transmission of corruption*

Esther Hauk
Universitat Pompeu Fabra†

Maria Saez Marti
Stockholm School of Economics‡

May 1999

Abstract

We provide a cultural explanation to the phenomenon of corruption in the framework of an overlapping generations model with intergenerational transmission of *values*. We show that the economy has two steady states with different levels of corruption. The driving force in the equilibrium selection process is the education effort exerted by parents which depends on the distribution of *ethics* in the population and on expectations about future policies. We propose some policy interventions which via parents' efforts have long-lasting effects on corruption and show the success of intensive education campaigns. Educating the young is a key element in reducing corruption successfully.

Keywords: dynamics of corruption, education, formation of preferences
JEL classification numbers: C73, K42, Z1

---

*This is an updated version of the working paper with the same title 98-30 of Universidad Carlos III de Madrid. We would like to thank Antonio Cabrales, Juan Carrillo, Marco Celentani, Albert Marcet, Joergen Weibull and Fabrizio Zilibotti for reading and commenting an earlier version of this paper. Esther Hauk gratefully acknowledges financial support of DGES under project PB96-0118. Maria Saez would like to thank the Comissionat per a Universitats i Recerca of the Generalitat de Catalunya for financial support and the European University Institute (Florence) for its hospitality.

†Corresponding author. Address: Universitat Pompeu Fabra, Department of Economics and Business, Ramon Trias Fargas 25-27, 08005 Barcelona, Spain. e-mail: esther.hauk@econ.upf.es

‡Box 6051, S-113 83 Stockholm, Sweden. e-mail: Maria.Saez-Marti@hhs.se

1 Introduction

> Mohammedans are Mohammedans because they are born and reared among the sect, not because they have thought it out and can furnish sound reasons for being Mohammedans; we know why Catholics are Catholics; why Presbyterians are Presbyterians; why Baptists are Baptists; why Mormons are Mormons; why thieves are thieves; why monarchists are monarchists; why Republicans are Republicans and Democrats, Democrats. We know that it is a matter of association and sympathy, not reasoning and examination; that hardly a man in the world has an opinion on morals, politics, or religion that he got otherwise than through his associations and sympathies. **Mark Twain**

Recent scandals in Japan, Italy and Spain show that corruption is not an exclusive phenomenon of underdeveloped countries. Countries with similar degrees of development exhibit enormous differences in the levels of corruption. Models linking corruption solely with the degree of economic development cannot explain these facts. Nor are they fully captured by institutional differences. Incentive and punishment schemes for corrupt activities certainly influence the level of corruption, but a given scheme does not work equally well in all countries. This suggests that corruption may be due, at least in part, to cultural elements.

Several papers have addressed the question why countries with the same level of development and similar institutions may have drastically different levels of corruption, using models with multiple steady states. In the existing literature, this multiplicity is reached by assuming some heterogeneity among economic agents, namely economic agents have heterogenous costs when engaging in corrupt activities (e.g. Andvig and Moene (1990), Cadot (1987), Carrillo (1998), Casagrande (1998a), Casagrande (1998b), Lui (1986), Sah (1988), Tirole (1996)). Typically, these costs are non-monetary or moral[1] and *exogenous* to the model. In our view, the assumption of a fixed exogenous structure of non-monetary costs is unsatisfactory: if we argue that different levels of corruption might be caused by different moral costs we need to explain why different countries may have different moral costs in the first place. The contribution of the present paper is to address this question. It focuses on the cultural transmission of corruption and develops policy measures which can reduce corruption within this framework. To keep the model simple, we assume that there are only two possible types of agents: *honest* "moral" agents who suffer some utility loss due to the feeling of guilt when engaging in corrupt activities and *potentially dishonest* agents who only care about monetary payoffs. Since morality is a result of socialization, the distribution of these types is endogenized in our model. This is done by modeling explicitly how

[1] The idea that morality can be incorporated into agents' preference ordering via feelings stretches back to Arrow(1967) and has been supported by a number of psychological, sociological and experimental studies (see e.g. Coursey et al. (1987) and Frank (1988)).

different attitudes against corruption are transmitted over different generations. The crucial point of this paper is that while these values are transmitted from one generation to the next, the incentives of parents to shape their offsprings' attitudes towards corruption depend on economic factors.

In this paper we focus on the intergenerational transmission of attitudes against corruption, incorporating both vertical transmission, with offsprings learning from their parents, and oblique transmission, with offsprings learning from some member of the parent's generation (see Boyd and Richerson (1985)). Cavalli-Sforza and Feldman (1982) provide evidence for vertical transmission of cultural traits such as religious beliefs, political attitudes, the frequency of praying and attending church, sportive practices, the frequency of listening to classical music, salt usage...etc. LeBras and Todd (1981) hypothesize that family structure strongly influences political beliefs. Phenomena like the Mafia seem to have their origin in societies where families are strong institutions and children are exposed, from the very beginning, to a homogenous set of cultural models in the family (Cavalli-Sforza (1996)).

An example of the influence of oblique transmission on corruption levels is Hong Kong where public attitudes against corruption changed drastically in the last decades due to the Independent Commission Against Corruption (ICAC) and especially its community relations department. The main emphasis of the ICAC education program was to "build a strong altruism and a sense of responsibility in oneself and toward the others", de-emphasizing the importance of getting money and getting ahead at the expenses of the others (Clark (1987))[2]. Some empirical studies point out that the perception of corruption as a social problem in Hong Kong depends to some extent on age (and therefore on the time the different groups were exposed to the ICAC). For instance, in 1986, 75.1% of the 15-24 age group (which had been subject to the ICAC's education program for about 13 years) believed that corruption was a social problem whereas only 54% of the 45-64 age group (who were born and lived their formative years when the ICAC didn't exist) agreed with that.

Hong Kong's anti-corruption measures are considered as very successful. Indeed many anti-corruption groups and governments encouraged an (academic) debate in order to assess whether and in how far to imitate Hong Kong. Examples are the Chicago Ethics Project who seriously considered implementing Hong Kong's policies[3], and China, whose State Administration of Industry and Commerce (SAIC) periodically launches ideological anti-corruption campaigns as a key part of all Chinese anti-

[2]The declared goals of ICAC were: "To change people's behaviour so that they will not engage in corrupt behaviour initially for fear of detection (deterrence), later because they cannot (prevention) and yet later because they do not wish to (attitude change). In order to achieve this, ICAC did not only rely on an education campaign but also changed the incentive system. These anti-corruption measures will be considered in our model.

[3]These considerations are reflected in a special issue of the journal "Corruption and Control" (1989) where academics like Gardiner, Clark and Johnston express their pros and contras to this idea.

corruption campaigns. In the present paper we examine under which circumstances education campaigns successfully reduce the level of corruption in the population. It will be shown that education campaigns do not necessarily reduce and might even increase corruption.

Vertical and oblique transmission of morality are incorporated into our theory by assuming that ethics against corruption are transmitted via education. The transmission model is similar to Bisin and Verdier (1996). A simple overlapping generation model with principal-agent relation, rational expectations and random matching is postulated. In each period, infinitely lived principals are randomly matched to the agents. At any time period an agent may give birth to a child who will become active in the next period. During the rearing period the parent has to educate his child. (Stochastic death keeps the population constant). Parents care about their children and want to maximize their child's well-being. However, given that they do not know what is best for their child, they evaluate their child's welfare as if it were their own[4]. Their own preferences are the best proxy they have for evaluating their child's well-being. Following Bisin and Verdier (1996) the cultural parent chooses the "coefficient of cultural transmission", or the education effort i.e., the probability with which the parent's cultural trait is adopted by the child. In the basic model, when the child does not "learn" from the parent, he imitates a randomly chosen member of the parent's generation[5]. In this world, if education were free, parents would choose to transmit their preferences with probability 1 and the society wouldn't evolve. If education in the family were prohibitively expensive, new agents would follow a "conformist" learning mechanism, and the spread of the most frequent trait would be observed. The higher the education effort, the smaller the importance of this *frequency-dependent bias* (Boyd and Richerson (1985)).

In our model, as in a typical principal-agent model, corruption exists because of asymmetric information and costly monitoring. In our information structure each principal knows the exact proportion of dishonest players in the population and has some (imperfect) information about the honesty of the agent he is facing. There is no information leakage across principals.

Our basic principal-agent model is related to Tirole (1996). In each period a principal has to assign a project to the agent he is randomly matched with. There are two types of projects. Project 1 is socially better than project 2 if managed with honesty. The reverse is true if the agent behaves dishonestly. The projects can be interpreted as two different public investments, one more costly than the other and with a higher social return if managed correctly. This project, by involving a larger amount of money, is more susceptible to corruption (selection of worse materials, manipulation of allocation mechanism such as auctions...). Think for instance of project 1 as the construction

[4]Alternatively, the model could be interpreted as parents caring about their child's behaviour or parents caring about the survival of their own preferences.

[5]In the extended model, moral education by the government is introduced.

of new roads and project 2 as resources devoted to the maintenance of existing ones.

We show that under reasonable parameters both types of agents choose positive transmission coefficients. This implies that any stable steady state is interior and that corruption is never eliminated completely as long as some corrupt behaviour existed in the past. We show that under rational expectations there are two pure strategy steady states. In the *low corruption steady state*, the principals offer project 1 to all agents for whom there is no evidence of corrupt behaviour and project 2 to all those agents for whom such evidence exists. In the *high corruption steady state* only project 2 is offered. In this case the existence of dishonest players exerts a negative externality on the honest players. The general suspicion prevents honest people from getting good projects like in Tirole (1996). However, while in Tirole's (1996) model cultural attitudes exist in fixed proportions, in the present paper the proportion of types evolves endogenously. Therefore, it is a result and not an assumption of the model that corruption is never eliminated completely in steady state and the effect of policies on the evolution of the proportions of different types can be analyzed. Three parameter regimes are distinguished: *(i)* the high corruption steady state is reached always, *(ii)* the low corruption steady state is reached always and *(iii)* one of the steady states is reached depending on initial conditions. *(i)* and *(ii)* result from extremely poor or nearly perfect monitoring technologies of principals. Therefore, it seems a reasonable guess to assume that most countries are likely to be in *(iii)*.

For the latter case the paper develops some temporary policy measures in order to permanently manipulate the cultural transmission coefficients. The advantages and disadvantages of each measure are discussed. Two time consistent policy measures with long-lasting effects on the level of corruption are proposed. The first policy consists of a temporary increase in the monitoring expenditure and, consequently, in the accuracy of the information gathered by the principals. Such a policy can drive the economy out of the high corruption steady state and into the basin of attraction of the low corruption equilibrium. In the second policy the principals announce a future policy change. This announcement triggers a change in the education efforts exerted by the different types of parents, which makes the policy announcement time consistent. We also discuss the effect of (temporary) public education campaigns and show that they successfully reduce the level of corruptions if and only if they are intensive enough, i.e. if the public education effort is high enough and the campaign is long-lasting. This condition seems to have been satisfied in the case of Hong Kong.

The paper is organized as follows. In section 2 we introduce the model and characterize the steady states. Policy implications are spelt out in section 3 and the effects of public education campaigns are discussed. Section 4 concludes.

2 The model

We propose a principal-agent model similar to Tirole's (1996). We consider a random matching model where each agent can never meet the same principal twice. At each time t ($-\infty < t < \infty$) every active agent is matched with a new principal. The principal gives the agent one of 2 projects. Project 1 yields a higher payoff to the principal than project 2 if the agent is honest, but is more susceptible to corrupt behaviour. The payoffs to the principal are

$$H > h \geq d > D$$

where capital letters denote the payoffs to the principal if project 1 is given. H stands for honest and D for dishonest behaviour by the agent.

Agents can be of two types: honest or potentially dishonest. The payoffs to an honest agent are as follows

honest type

	Project 1	Project 2
honest	B	b
dishonest	$\bar{B} - e$	$\bar{b} - e$

With $B, \bar{B}, b, \bar{b}, e > 0$, $B > b$, $\bar{B} > \bar{b}$ and

$$e > \bar{B} - B \geq \bar{b} - b \geq 0. \tag{1}$$

If (1) holds, honest agents always behave honestly. Observe that an honest agent suffers from being dishonest. He is endowed with a *moral attitude* which favours "honest" behaviour. On the contrary, potentially dishonest agents only care about monetary payoffs,

potentially dishonest type

	Project 1	Project 2
honest	B	b
dishonest	$\bar{B}$	$\bar{b}$

Under (1) potentially dishonest agents are always dishonest. Thereafter, since we assume that (1) holds, we shall refer to potentially dishonest players as dishonest.

The model is a model of overlapping generations. A Poisson birth and death process is assumed keeping the population size of active agents constant. With probability λ an active agent will be active next period. With probability $(1-\lambda)$ an active agent in t has a child which at the moment of birth does not have any predetermined preferences. The child is educated by the parent and becomes active in $t+1$. Education shapes the child's preferences. The crucial assumption is that an agent cares about his offsprings' welfare and tries to maximize the latter when deciding how much effort to put into his child's education. Given that at the moment of education, the new born does not have any preferences, the parent evaluates his child's future utility through his own eyes. In other words he uses his payoff matrix as if it were his child's, like in Bisin and Verdier (1996).

The education process works as follows: The parent educates his naive child with some education effort τ. With probability equal to the education effort, education will be successful and the child will be like his parent[6]. Otherwise, the child remains naive and gets randomly matched with somebody else whose preferences he will adopt. Consider an honest agent who has a child at time t and chooses education effort τ^a and let p_t^{ij} be the probability that a child of parent i is of type j

$$p_t^{aa} = \tau_t^a + (1-\tau_t^a)q_t \quad (2)$$

$$p_t^{ab} = (1-\tau_t^a)(1-q_t) \quad (3)$$

where q_t is the proportion of honest agents at time t. Similarly, for the dishonest parent we get

$$p_t^{bb} = \tau_t^b + (1-\tau_t^b)(1-q_t) \quad (4)$$

$$p_t^{ba} = (1-\tau_t^b)q_t \quad (5)$$

where τ^b is the dishonest parents' education effort.

2.1 The education choice

Let $C(\tau)$ be the cost of the education effort τ and assume that $C(0)=0$, $C'>0$ and $C''>0$.

A parent of type i chooses the education effort $\tau \in [0,1]$ that maximizes

$$p_t^{ii}V_t^{ii} + p_t^{ij}V_t^{ij} - C(\tau) \quad (6)$$

where p^{ij} and p^{ii} are defined above and V^{ij} is the utility a parent with preferences i attributes to his child having preferences j. In order to assess V^{ij} a parent of type i uses his own payoff matrix. Therefore $V^{ii} > V^{ij}$ always. Notice that V^{ij} is an expected utility and depends on the policy expectations of the parent. Maximizing (6) with respect to τ we get the following first order condition

$$C'(\tau^i) = \frac{dp^{ii}}{d\tau^i}V^{ii} + \frac{dp^{ij}}{d\tau^i}V^{ij} \quad (7)$$

[6] Parents believe it is optimal for their child to be like themselves: while honest parents believe that their child will suffer if behaving dishonestly, dishonest parents believe that their child will not maximize its monetary returns if honest. These are "reasonable" believes: both types of parents behave optimally given their preferences.

where we have suppressed the time indicators.

Substituting (2)-(5) in (7), we get the optimal education efforts τ^a and τ^b,

$$C'(\tau^a) = (V^{aa} - V^{ab})(1-q) \tag{8}$$

$$C'(\tau^b) = (V^{bb} - V^{ba})q \tag{9}$$

In order to have interior solutions $\tau \in (0,1)$ we need that $C'(0) = 0$ and that $C'(1) > \bar{B}/(1-\lambda)$, which is the upper bound to agents' payoffs. From (8) and (9) it follows that the optimal effort level is $\tau^i = \tau(q, V^{ii} - V^{ij})$ with

$$\frac{\partial \tau^a(q, V^{aa} - V^{ab})}{\partial q} = -\frac{V^{aa} - V^{ab}}{C''(\tau^a(q, V^{aa} - V^{ab}))} < 0 \quad \text{and}$$

$$\frac{\partial \tau^b(q, V^{bb} - V^{ba})}{\partial q} = \frac{V^{bb} - V^{ba}}{C''(\tau^b(q, V^{bb} - V^{ba}))} > 0.$$

Since $V^{ii} - V^{ij}$ depends on the parent's policy expectations, so does the optimal effort level $\tau^i(q, V^{ii} - V^{ij})$.

We can now characterize the dynamic behaviour of q_t:

$$q_{t+1} = \lambda q_t + (1-\lambda)(q_t p_t^{aa} + (1-q_t)p_t^{ba})$$

substituting (2) and (5), we obtain

$$\begin{aligned} q_{t+1} &= F(q_t, V^{aa} - V^{ab}, V^{bb} - V^{ba}) = \\ & q_t + (1-\lambda)q_t(1-q_t)(\tau^a(q_t, V^{aa} - V^{ab}) - \tau^b(q_t, V^{bb} - V^{ba})) \end{aligned}$$

which can be rewritten (suppressing the time indices) as

$$F(q, V^{aa} - V^{ab}, V^{bb} - V^{ba}) - q = (1-\lambda)q(1-q)(\tau^a(q, V^{aa} - V^{ab}) - \tau^b(q, V^{bb} - V^{ba})) \tag{10}$$

Observe that (10) has three rest points: *i)* $q = 0$, *ii)* $q = 1$ and *iii)* $q = q^*$,

$$q^* = \frac{V^{aa} - V^{ab}}{V^{bb} - V^{ba} + V^{aa} - V^{ab}} \tag{11}$$

with $\tau^a(q^*, V^{aa} - V^{ab}) = \tau^b(q^*, V^{bb} - V^{ba})$.

Lemma 1 *Assume that $C''(\tau) \geq C'(\tau) > 0$ for all $\tau \neq 0, C'(0) = 0$ and $C'(1) > \max_i V^{ii} - V^{ij}$. The rest points 0 and 1 are unstable and q^* is globally stable.*

Proof. See appendix.

2.2 The principals' choice

Each period a principal has to decide what project to delegate on the agent he is matched with. We assume that principals maximize their expected payoffs and that they know the proportion of honest agents in the population but not the type of a particular agent. We assume that the principal can know with positive probability α wether the agent he is facing is dishonest[7]. An honest agent will never be revealed as dishonest. There is no information leakage across principals[8]. If one principal learns that an agent is dishonest it can still be the case that in the future the same agent is taken for an honest one.

Let σ^s be the *separating strategy* consisting of offering project 1 to seemingly honest agent and project 2 to agents who are found dishonest. Assume that principals follow strategy σ^s, then potentially dishonest player will behave dishonestly if

$$B < (1-\alpha)\bar{B} + \alpha\bar{b}$$

which can be rewritten as

$$\alpha < \frac{\bar{B} - B}{\bar{B} - \bar{b}}. \tag{12}$$

Thereafter we will assume that (12) holds.

Let σ^p be the *pooling strategy* of offering project 2 to everybody. Principals prefer strategy σ^s to σ^p if

$$q_t(H-h) + (1-q_t)(1-\alpha)(D-d) > 0 \tag{13}$$

which can be rewritten as

$$q_t > \frac{(1-\alpha)(d-D)}{(H-h)+(d-D)(1-\alpha)} \equiv \tilde{q}. \tag{14}$$

Let $\sigma(q_t)$ denote the principals' optimal strategy at time t, namely

$$\sigma(q_t) = \begin{cases} \sigma^s & if \quad q_t > \tilde{q} \\ \{\sigma^s, \sigma^p\} & if \quad q_t = \tilde{q} \\ \sigma^p & if \quad q_t < \tilde{q} \end{cases}$$

[7]Tirole (1996) assumes that the principal has some imperfect information about each agent's past behaviour: with probability α he knows if the agent has been dishonest at least once in the past. Under this information structure corrupt new borns are indistinguisable from honest agents. With Tirole's story the qualitative results are the same but calculations are much more cumbersome. Notice that in Tirole the gain from being corrupt is higher in some cases since in the first period of a dishonest's life cheating cannot be detected.

[8]Information leakage across principals does not affect the qualitative results of the paper.

2.3 The steady states.

We now characterize the steady states of the economy. The education effort exerted by a parent in t depends on the expectation about the principals' policy in the future. A "policy" is an (infinite) sequence $\{\sigma_z\}_{z=t_1}^{\infty}$, with $\sigma_z \in \{\sigma^s, \sigma^p\}$, for all z. We will denote by $\{\sigma^i\}_{t_1}^{t_2}$, the sequence consisting of the repetition of σ^i from t_1 to t_2 ($t_1 < t_2 \leq \infty$). Let $V^{ij}(k_t^e)$ be the expected utility a parent of type i attributes to his child born in t (and active in $t+1$) having preferences j when the expected policy is k_t^e and let $\tau^i(q_t, k_t^e) = \tau^i(q_t, V^{ii}(k_t^e) - V^{ij}(k_t^e))$ be the education effort of a parent of type i in t who expects a policy $k_t^e = \{\sigma_z\}_{z=t+1}^{\infty}$.

Lemma 2 *Assume $C'(\tau) > 0$ and that condition (12) holds. Then*

1. $\tau^a(q_t, \{\sigma^s\}_{t+1}^{\infty}) \gtreqless \tau^b(q_t, \{\sigma^s\}_{t+1}^{\infty})$, when $q_t \lesseqgtr \bar{q}$
2. $\tau^a(q_t, \{\sigma^p\}_{t+1}^{\infty}) \gtreqless \tau^b(q_t, \{\sigma^p\}_{t+1}^{\infty})$, when $q_t \lesseqgtr \underline{q}$
3. $\tau^a(q_t, \{\{\sigma^s\}_{t+1}^{T-1}, \{\sigma^p\}_{T}^{\infty}\}) \gtreqless \tau^b(q_t, \{\{\sigma^s\}_{t+1}^{T-1}, \{\sigma^p\}_{T}^{\infty}\})$, when $q_t \lesseqgtr \bar{q} - \lambda^{T-t-1}(\bar{q} - \underline{q})$,
4. $\tau^a(q_t, \{\{\sigma^p\}_{t+1}^{T-1}, \{\sigma^s\}_{T}^{\infty}\}) \gtreqless \tau^b(q_t, \{\{\sigma^p\}_{t+1}^{T-1}, \{\sigma^s\}_{T}^{\infty}\})$, when $q_t \lesseqgtr \underline{q} + \lambda^{T-t-1}(\bar{q} - \underline{q})$, where

$$\underline{q} = \frac{e - (\bar{b} - b)}{e} \quad \text{and} \quad \bar{q} = \frac{e + \alpha(\bar{B} - \bar{b}) - (\bar{B} - B)}{e}.$$

Proof. From (8)-(9) we get that $\tau^a(q_t, k_t^e) > \tau^b(q_t, k_t^e)$ when

$$q_t < \frac{V^{aa}(k_t^e) - V^{ab}(k_t^e)}{V^{bb}(k_t^e) - V^{ba}(k_t^e) + V^{aa}(k_t^e) - V^{ab}(k_t^e)} \tag{15}$$

Computing the right hand side of (15) for the different expected policy profiles we get the values above. □

The previous lemma compares the education efforts exerted by the two types of parent for four different expectations, two of them stationary (cases 1 and 2) and two of them involving a policy change at a future date T (cases 3 and 4). Observe that $\bar{q} > \underline{q}$ when

$$\alpha > \frac{(\bar{B} - B) - (\bar{b} - b)}{\bar{B} - \bar{b}} \tag{16}$$

Proposition 1 *Assume $C''(\tau) \geq C'(\tau) > 0$ for all τ, $q_0 \neq \{0, 1\}$, (12), (16) hold, principals follow $\sigma(q_t)$ and agents have rational expectations. Then,*

1. *q_t converges to $\bar{q}$ if $\tilde{q} < \underline{q}$,*

2. *q_t converges to $\underline{q}$ if $\tilde{q} > \bar{q}$ and*

3. *when $\underline{q} < \tilde{q} < \bar{q}$*

    (a) *q_t converges to $\bar{q}$ if $q_0 > \tilde{q}$ and*

    (b) *q_t converges $\underline{q}$ if $q_0 < \tilde{q}$.*

Proof. See appendix .

We refer to $\bar{q}$ and $\underline{q}$ as the *low corruption* and the *high corruption* steady states, respectively.

Observe that the steady state the system converges to depends for cases 1 and 2 on the location of $\tilde{q}$ with respect to $\bar{q}$ and $\underline{q}$ and in case 3, on the initial proportion of honest players. A too inefficient monitoring technique (high $\tilde{q}$) implies that the low corruption steady state $\bar{q}$ can never be reached. A very efficient technique avoids the high corruption steady state. The smaller $\tilde{q}$ the larger the basin of attraction of $\bar{q}$.

3 Policy measures.

Under rational expectations the steady state the system converges to is determined by the relative positions of $\bar{q}$, $\underline{q}$ and $\tilde{q}$ and in the case in which $\underline{q} < \tilde{q} < \bar{q}$ also by the initial proportion of honest agents. While the position of $\underline{q}$ only depends on the payoff matrices of the agents, $\bar{q}$ and $\tilde{q}$ also depend on the accuracy of the principals' information α. Hence, feasible policy measures will have to affect the remuneration to agents or the accuracy of principals' information or agents' expectations. We shall now discuss the advantages and disadvantages of these measures.

Changing the remuneration to the agents will affect equilibrium values directly. An increase in the payoff when agents behave honestly in project 1 (B) and in project 2 (b) increases the equilibrium proportion of honest players in the low and in the high corruption equilibria, respectively. The same is true for a decrease of $\bar{B}$ and $\bar{b}$. However, principals will face some restrictions when choosing remunerations. Higher wages simultaneously increase B (or b) and $\bar{B}$ (or $\bar{b}$). In order to lower $\bar{B}$ or $\bar{b}$ principals would have to be able to limit the extent of corrupt activities somehow.

Another possibility to control for corruption is to invest in monitoring. Notice, that an increase in the accuracy of principals' information α and thereby in the probability of detecting fraudulent behaviour will shift the critical value $\tilde{q}$ to a lower value, will increase the basin of attraction of the low corruption steady state and at the same time the proportion of honest behaviour in such an equilibrium. Since the choice of α is influenced by technological restrictions, unless some new monitoring technology is discovered, it is not reasonable to assume that principals can improve their information forever. The following argument shows that a temporary increase in spending on monitoring might be sufficient to leave the high corruption steady state.

Assume that $\tilde{q} < \bar{q}$ and that there exists an $\bar{\alpha} > \alpha$ such that $\tilde{q}(\bar{\alpha}) = \underline{q}$, then the high corruption steady state can be left by a temporary increase in α. Given $\bar{\alpha}$, the separating strategy σ^s is optimal and seemingly honest people will get project 1. By lemma 2 the high corruption steady state will be left if honest agents expect $k_t^e = \{\sigma^s\}_{t+1}^{\infty}$. The principal can ensure this by reducing spending on monitoring to the original level in such a way that $\tilde{q}(\alpha_t)$ is always smaller than q_t in all periods t.

In the above policy measure principals behave optimally given the accuracy of their information α. Therefore, if agents can observe how much principals spent on monitoring the policy measure is perfectly credible. Moreover, this policy is feasible if the temporary increase in spending is off-set by the gains from reaching the low corruption steady state.

By a similar argument, principals could incur a cost by simply giving the good project in a bad environment ($q_t < \tilde{q}$) to stimulate education efforts of honest parents. In other words principals would have to apply the separating policy σ^s despite its being sub-optimal in the short run. For this policy to be effective agents would have to believe that principals are willing to ignore their cut-off value over several periods. In contrast, a temporary increase in spending on monitoring differs from simply ignoring the cut-off value since, under the former, the resulting behaviour of principals is always optimal given the observed increase in monitoring costs.

All policy measures discussed so far affected agents' expectations indirectly. We now consider an alternative policy which affects agents' expectations directly: the high corruption steady state can be left by principals announcing a time consistent policy change in the future. Under this policy measure, principals will never ignore their cut-off value and therefore behave optimally both in the short and long run.

Assume that the economy is in the high corruption steady state; everybody is getting project 2. In the high corruption steady state no principal has an incentive to give project 1 to anybody. Assume now, that at t principals commit to the policy profile $\{\{\sigma^p\}_{t+1}^{T-1}, \{\sigma^s\}_T^{\infty}\}$, namely, they will offer project 2 to everybody (pooling strategy) until time $T-1$, and from T onwards project 1 will be offered only to those seemingly honest agents (separating strategy).

Proposition 2 *Assume that* $q_t = \underline{q}$. *Policy* $\{\{\sigma^p\}_{t+1}^{T-1}, \{\sigma^s\}_T^{\infty}\}$ *is credible if* $q_{T-1} \leq \tilde{q} \leq q_T$

Proof. See appendix

[Include figure 1 here]

Figure 1 is an illustration of proposition 2. The economy is initially in the high corruption equilibrium. The continuous line is the value of $\tilde{q}$. The announcement of the policy change in $T = 15$ increases the honest parents' effort today and q starts growing. At $T = 14$ the economy has reached the critical value $\tilde{q}$ and from $T = 15$ on it is optimal to start offering project 1 to seemingly honest agents. This way the society can leave the high corruption equilibrium after the time consistent announcement. The policy announcement is time consistent because at the moment of the change the proportion of honest players in the population is such that (14) is satisfied and σ^s is the optimal policy sequence from then onwards. Observe that by proposition 1, the system converges to the low corruption steady state. The result is driven by the fact that the policy announcement raises the value of being honest more than it increases the value of being dishonest. Honest parents, foreseeing that in the future it will pay to be honest increase their education effort once such a policy is announced.

The above policy measures all aim at affecting the driving force of the population dynamics: the education effort exerted by parents. In the above model, moral education is a purely private issue. We shall now analyze the effectiveness of education campaigns in which the existing public education systems are used to emphasize moral values.

In most countries public education does not start immediately when a child is born. Usually, children are exposed to the influence of their parents before undergoing public education. To respect this common education structure we assume that only children who remain naive, i.e. who do not learn their preferences from their parents, can be influenced by public education. An education campaign will be modeled as society (or principals) investing in public moral education by choosing a public effort level ρ to teach honest behaviour at school. Similar to private education efforts, the public education effort represents the probability with which a child who did not learn from his parents adopts honest preferences in school. The new timing of moral education is as follows: as before, the education effort of the parents τ determines the probability with which children adopt the same preferences as their parents. With the complementary probability $(1 - \tau)$ children remain naive in which case the public education effort ρ determines the probability with which children become honest. With probability $(1 - \tau)$ public education fails and children meet a random member of society whose preferences they adopt.

Public education affects the probabilities of honest and dishonest children as follows[9]:

$$
\begin{aligned}
p_t^{aa} &= \tau_t^a + (1 - \tau_t^a)(q_t(1 - \rho) + \rho) && (17)\\
p_t^{ab} &= (1 - \tau_t^a)(1 - q_t)(1 - \rho) && (18)\\
p_t^{bb} &= \tau_t^b + (1 - \tau_t^b)(1 - q_t)(1 - \rho) && (19)\\
p_t^{ba} &= (1 - \tau_t^b)(q_t(1 - \rho) + \rho) && (20)
\end{aligned}
$$

[9] $\rho = 0$ is identical to the case without public education

The first order conditions which determine the private education efforts are now,

$$C'(\tau^a) = [V^{aa} - V^{ab}](1-q)(1-\rho) \tag{21}$$

$$C'(\tau^b) = [V^{bb} - V^{ba}](q(1-\rho)+\rho) \tag{22}$$

The new population dynamics are given by the following difference equation for q_t:

$$\begin{aligned} \Delta q \;=\; & (1-\lambda)q(1-q)(\tau^a(q, V^{aa} - V^{ab}) - \tau^b(q, V^{bb} - V^{ba}))(1-\rho) + \\ & (1-\lambda)(1-q)(1-\tau^b(q, V^{bb} - V^{ba}))\rho \end{aligned} \tag{23}$$

which can be rewritten as

$$\Delta q = (1-\lambda)(1-q)\left[(\tau^a(q, V^{aa} - V^{ab}) - \tau^b(q, V^{bb} - V^{ba}))q(1-\rho) + (1-\tau^b(q, V^{bb} - V^{ba}))\rho\right] \tag{24}$$

This difference equation shows that (i) $q = 1$ is always a rest point of the system, (ii) $q = 0$ is only a rest if no public education exists ($\rho = 0$), (iii) if an interior solution exists, the education effort of dishonest parents is higher than of honest parents ($\tau^a < \tau^b$). The introduction of public education has two opposite effects: while its direct effect is to increase the proportion of honest agents, its indirect effect is to change the incentives for private education; honest parents educate less because public education increases the chances of their children getting the right preferences anyway while dishonest parents educate more. Which effect will dominate partly depends on the value of ρ. Notice, that if ρ=1 the system converges to $q = 1$ although honest parents do not educate their children at all. Hence, for $\rho = 1$, $\Delta q > 0$ for all $q < 1$. By continuity, there exists a $\bar{\rho}$ such that for $\rho > \bar{\rho}$ $\Delta q > 0$ for all $q < 1$. Indeed, it is easy to see that for $\rho > \tau^b(1)$ $q = 1$ is the only attractor[10].

The above analysis establishes the success of a temporary intensive education campaign with a high enough ρ. Suppose society is in the high corruption steady state and $\tilde{q} < \bar{q}$. The government launches a very intensive education campaign with $\rho > \tau^b(1)$. The campaign affects the population dynamics and the proportion of honest agents increases. The education campaign can be stopped once $q_t > \tilde{q}$; by lemma 2 the system converges to the low corruption steady state $\bar{q}$.

[10]This is not the cut-off value. A complete analysis of the model becomes very messy and is beyond the scope of the paper. We are only interested in finding some temporary education campaign which is successful.

To summarize, temporary education campaigns will be successful in reducing corruption if they are intensive enough. If ρ is too low, q_t will remain below $\tilde{q}$ and it will not be optimal for principals to switch to the separating policy. Education campaigns work only if the investment in public education is high enough during the period of the campaign and the campaign lasts long enough.

Both conditions seemed to have been satisfied in the case of Hong Kong. The education effort of the Independent Commission against Corruption (ICAC) had been very high and the project lasted a substantial period of time. Moreover, at least in early years, ICAC combined two policy measures discussed in our model: re-education and a change in the remunerations to agents to reduce the profitability of corruption. This combination accelerates the move towards the low corruption steady state.

4 Conclusion

There is evidence that corruption is at least partly due to cultural elements. Not in every country does the public opinion conceive corruption - at least small-scale corruption - as negative. Sentences like "I was corrupt but so was everybody else" reveal that a generally corrupt environment can serve as a justification for one's own corrupt behaviour.

The present paper captures some cultural aspects of corruption. An agent is corrupt if corruption maximizes utility. However, utility is not only affected by purely monetary rewards but also by the presence (or absence) of moral costs if engaging in corrupt activities. In the model remunerations were chosen such that an agent is either always honest or always corrupt. Analyzing this extreme case allows to single out the purely educational effects on corruption levels. In order to do so it was assumed that new-born agents had to form their preferences and were influenced by the education effort exerted by their parent as well as by the general corruption level of society. Parents care about their children and judge their children's utility as if it were their own. The resulting dynamics had the realistic feature that the lower the proportion of a given type the higher its education effort. This feature keeps the steady state off the boundary and avoids a complete elimination of corrupt (or honest) agents.

Taking the model seriously implies that corruption will never be eliminated completely, a view which is also expressed by Klitgaard (1988). Indeed, there is no country without corruption although corruption levels vary widely across countries even with similar economic characteristics. The present model found two steady states one with a low and one with a high level of corruption in an otherwise identical economy. This shows the strength of cultural elements in determining the actual corruption levels of a society and implies that two countries with the same level of development and the same institutions against corruption may have drastically different levels of corruption depending on the initial state of the society. In the high corruption steady state the public reputation outweighs individual reputation and thereby locks society into highly corrupt behaviour. In the low corruption steady state individual reputation is decisive.

Controlling corruption imposes a cost on society. Individual behaviour has to be monitored. If monitoring is common and the technique is reliable, it pays less to be corrupt. This is also true for high fines. Both the present model and models concentrating purely on monetary rewards share this desirable feature. The advantage of the present approach is that it entails additional policy implications which can be cost-saving in the long run. High fines and high monitoring work only as long as they are implemented. If, however, young generations are educated to adapt a moral attitude against corruption, monitoring can be reduced while low corruption levels are preserved. Educating the young is the key element in reducing corruption successfully.

References

Akerlof, G. (1983). "Loyalty Filters," *American Economic Review*, **73**, 54–63.

Andvig, J.C. and K. Moene (1990). "How Corruption May Corrupt," *Journal of Economic Behavior and Organization*, **13**, 63–76.

Arrow, K.J (1984) "The place of Moral Obligations in Preference Systems", in Arrow *Collected Papers of Kenneth J. Arrow*, Blackwell

Besley, T. and McLaren, J. (1993). "Taxes and Bribery: The Role of Wage Incentives," *Economic Journal*, **103**, 119–141.

Bisin, and Verdier (1996). "The Economics of Cultural Transmission and the Dynamics of preferences," mimeo, MIT.

Boyd, R. and Richerson, P. (1985). *Culture and the Evolutionary Process*, University of Chicago Press, Chicago

Block, M.K. and Heineke, J.M. (1975). "A Labour Theoretical Analysis of Criminal Choice," *American Economic Review*, **65**, 314–325.

Cadot, O. (1987) "Corruption as a Gamble", *Journal of Political Economics* 33, 223-244

Carrillo, D. J., (1998). "Grafts, Bribes and the Practice of Corruption", mimeo, ECARE and CEPR

Casagrande, A. (1998a), "Tax Evasion and a role for the government: a general equilibrium approach", mimeo, New York University

Casagrande, A. (1998b), "Calibrating tax evasion", mimeo, New York University

Cavalli-Sforza, L. (1996). *Geni, popoli e lingue.* Adelphi Edizioni, Milano

Cavalli-Sforza, L. and Feldman, M. (1981). *Cultural Transmission and Evolution.* Princton University Press, Princton.

Chinn, D. (1978). "Farmer Response to Good Grain Controls in Developing Countries," *Quarterly Journal of Economics* , **92**, 697–703.

Clark, D. (1987). "A community relations approach to corruption: the case of Hong Kong," *Corruption and Reform:*, **2**, 135–257

Clark, D. (1989). "Mobilizing public opinion against corruption: a commentary," *Corruption and Reform:*, **4**, 123–129.

Coursey D.L, E. Hoffman and M. Spitzer (1987). "Fear And Loathing in the Coase Theorem: Experimental Evidence Involving Physical Discomfort" *Journal of Legal Studies*, **16**, 217–248.

Frank, R.H. (1988). *Passion Within Reason: The Strategic Role of Emotions.* New York: Norton and Company

Klitgaard, R. (1988). *Controlling corruption.* University of California Press, Berkeley.

Laffont, J. and T. N'Guessan (1998). "Competition and Corruption in an Agency Relationship", mimeo, Toulouse.

LeBras, H. and E. Todd. (1981). *L'invention de la France: Atlas Anthropologique et Politique,* Livre de Poche, Hachette, Paris.

Lui, F.T. (1986), "A dynamic model of corruption deterrence", *Journal of Public Economy* **31**, 215-236

Manion, M. (1996). "Corruption by Design: Bribery in Chinese Enterprise Licensing," *Journal of Law, Economics and Organization*, **12**, 167-195.

Qizilbash, Mozaffar (1994). "Corruption, Temptation and Guilt: Moral Character in Economic Theory," Discussion Paper No. 9419, University of Southampton.

Rose-Ackerman, S. (1978). *Corruption: A Study in Political Economy.* Academic Press

Sah, R.K. (1988). "Persistence and pervasiveness of corruption: new perspectives", Yale University Economic Growth Center Discussion Paper 560.

TIROLE, J. (1996). "A theory of collective reputation (with applications to the persistence of corruption and firm quality," *Review of Economic Studies*, **63**, 1–22.

APPENDIX

Proof of Lemma 1

To simplify the notation we eliminate $V^{ii} - V^{ij}$ from the arguments in the functions $F(\cdot)$ and $\tau(\cdot)$.

$$F(q)=q+(1-\lambda)q(1-q)(\tau^a(q)-\tau^b(q))$$

Observe that $\tau^a(q) > \tau^b(q)$ for all $q \in [0, q^*)$ and $\tau^a(q) < \tau^b(q)$ for all $q \in (q^*, 1]$. Therefore $F(q) > q$ when $q \in [0, q^*)$ and $F(q) < q$ when $q \in (q^*, 1)$.

$$F'(q) = 1 + (1-\lambda)(1-2q)(\tau^a(q)-\tau^b(q)) - (1-\lambda)q(1-q)[\frac{V^{aa}-V^{ab}}{C''(\tau^a(q))} + \frac{V^{bb}-V^{ba}}{C''(\tau^b(q))}] \quad (25)$$

substituting (8) and (9) in (25)

$$F'(q) = 1 + (1-\lambda)(1-2q)(\tau^a(q)-\tau^b(q)) - (1-\lambda)[\frac{qC'(\tau^a(q))}{C''(\tau^a(q))} + \frac{(1-q)C'(\tau^b(q))}{C''(\tau^b(q))}]$$

Evaluating $F'(q)$ in q^* we get

$$F'(q^*) = 1 - (1-\lambda)\frac{C'(\tau^*)}{C''(\tau^*)} \in (0,1)$$

where $\tau^* = \tau^a(q^*, V^{aa} - V^{ab}) = \tau^b(q^*, V^{bb} - V^{ba})$. This shows that q^* is locally stable. Since $F'(0) > 1$ and $F'(1) > 1$, 0 and 1 are unstable. Moreover $F'(q) > 0$ for all $q \in [0, \min\{1/2, q^*\}] \cup [\max\{1/2, q^*\}, 1]$. This, together with the fact that $F(q) > q$ when $q \in (0, q^*)$ and $F(q) < q$ when $q \in (q^*, 1)$ implies that q_t converges to q^* for all $q_t \in (0,1)$. Hence q^* is globally stable. □

Remark. Observe that we cannot characterize $F'(q)$ in the interval $(\min\{1/2, q^*\}, \max\{1/2, q^*\})$ without further assumptions on the cost function $C()$,although it is easy to see that

$$F'(q) > 1 - (1-\lambda) \mid 1 - 2q^* \mid -(1-\lambda) \quad (26)$$

The right hand side in (26) is positive when

$$\lambda > 1 - \frac{1}{1+ \mid 1-2q^* \mid} \in [0, \frac{1}{2}]$$

If $\lambda > 1/2$, $F'(q) > 0$ for all q.

Proof of Proposition 1.

Case 1: $\tilde{q} < \underline{q}$

1.a) Consider the expected policy profile $\{\sigma^s\}_{t+1}^{\infty}$. By lemma 2, $\tau^a(q_t, \{\sigma^s\}_{t+1}^{\infty}) \gtreqless \tau^b(q_t, \{\sigma^s\}_{t+1}^{\infty})$ for all $q_t \lesseqgtr \bar{q}$.

$$V_t^{aa}(\{\sigma^s\}_{t+1}^{\infty}) - V_t^{ab}(\{\sigma^s\}_{t+1}^{\infty}) = \frac{e\bar{q}}{1-\lambda} > 0 \tag{27}$$

$$V_t^{bb}(\{\sigma^s\}_{t+1}^{\infty}) - V_t^{ba}(\{\sigma^s\}_{t+1}^{\infty}) = \frac{e(1-\bar{q})}{1-\lambda} > 0 \tag{28}$$

Given $\{\sigma^s\}_{t+1}^{\infty}$, $\bar{q}$ is globally stable (by lemma 1). For all $q > \tilde{q}$, $\sigma(q) = \sigma^s$ and $\sigma(q_{t+1})_{t \geq 0} = \{\sigma^s\}_{t+1}^{\infty}$ if $q_t > \tilde{q}$.

1.b) Consider the expected policy profile $\{\sigma^p\}_{t+1}^{\infty}$. By lemma 2, $\tau^a(q_t, \{\sigma^p\}_{t+1}^{\infty}) \gtreqless \tau^b(q_t, \{\sigma^p\}_{t+1}^{\infty})$ for all $q_t \lesseqgtr \underline{q}$.

$$V^{aa}(\{\sigma^p\}_{t+1}^{\infty}\}) - V^{ab}(\{\sigma^p\}_{t+1}^{\infty}\}) = \frac{e\underline{q}}{1-\lambda} > 0 \tag{29}$$

$$V^{bb}(\{\sigma^p\}_{t+1}^{\infty}\}) - V^{ba}(\{\sigma^p\}_{t+1}^{\infty}\}) = \frac{e(1-\underline{q})}{1-\lambda} > 0 \tag{30}$$

Given $\{\sigma^p\}_{t+1}^{\infty}$, $\underline{q}$ is globally stable (by lemma 1). We can find a $t > 0$ such that q_t is arbitrarily close to $\underline{q}$.

1.c) Assume now that $q_t < \tilde{q} < \underline{q}$ and consider the expected policy profile $\{\{\sigma^p\}_{t+1}^{\mathrm{T}-1}, \{\sigma^s\}_{\mathrm{T}}^{\infty}\}$

$$V_t^{aa}(\{\{\sigma^p\}_{t+1}^{T-1}, \{\sigma^s\}_T^{\infty}\}) - V_t^{ab}(\{\{\sigma^p\}_{t+1}^{T-1}, \{\sigma^s\}_T^{\infty}\}) = \frac{e((1-\lambda^{\mathrm{T}-t-1})\underline{q} + \lambda^{\mathrm{T}-t-1}\bar{q})}{1-\lambda} > 0, \tag{31}$$

$$V_t^{bb}(\{\{\sigma^p\}_{t+1}^{T-1}, \{\sigma^s\}_T^{\infty}\}) - V_t^{ba}(\{\{\sigma^p\}_{t+1}^{T-1}, \{\sigma^s\}_T^{\infty}\}) = \frac{e((1-\lambda^{\mathrm{T}-t-1})(1-\underline{q}) + \lambda^{\mathrm{T}-t-1}(1-\bar{q}))}{1-\lambda} > 0, \tag{32}$$

Observe that (31) is decreasing and (32) is increasing in T. This implies that

$$\tau^a(q_t, \{\{\sigma^p\}_{t+1}^{T-1}, \{\sigma^s\}_T^{\infty}\}) - \tau^b(q_t, \{\{\sigma^p\}_{t+1}^{T-1}, \{\sigma^s\}_T^{\infty}\})$$

is decreasing in T for all q. For the same initial condition $q_0 < \underline{q}$, q_t is larger the smaller is T for all $t > 0$ and q_T is larger the larger is T. Notice that

$$\tau^a(q_t, \{\{\sigma^p\}_{t+1}^{T-1}, \{\sigma^s\}_T^\infty\}) - \tau^b(q_t, \{\{\sigma^p\}_{t+1}^{T-1}, \{\sigma^s\}_T^\infty\}) > \tau^a(q_t, \{\sigma^p\}_{t+1}^\infty) > \tau^b(q_t, \{\sigma^p\}_{t+1}^\infty)$$

for all T. There exist a finite $\hat{T}$ such $q_{\hat{T}-1} \leq \tilde{q}$, and $F(q_{\hat{T}-1}) \geq \tilde{q}$, since under $\{\sigma^p\}_{t+1}^\infty$, $\tilde{q}$ is reached in finite time.

>From 1.a) and 1.c) we conclude that $\{\{\sigma^p\}_{t+1}^{\hat{T}-1}, \{\sigma^s\}_{\hat{T}}^\infty\} = \sigma(q_{t+1})_{t\geq 0}$, and q_t converges to $\bar{q}$.

Case 2: $\tilde{q} > \bar{q}$

2.a) If $q_t < \tilde{q}$, $\{\sigma^p\}_{t+1}^\infty = \sigma(q_{t+1})_{t\geq 0}$ and q_t converges to $\underline{q}$. (see part b) above).

2.b) $q_t > \tilde{q}$. Consider the policy $\{\{\sigma^s\}_{t+1}^{T-1}, \{\sigma^p\}_T^\infty\}$,

$$V_t^{aa}(\{\{\sigma^s\}_{t+1}^{T-1}, \{\sigma^p\}_T^\infty\}) - V_t^{ab}(\{\{\sigma^s\}_{t+1}^{T-1}, \{\sigma^p\}_T^\infty\}) = \frac{e((1-\lambda^{T-t-1})\bar{q} + \lambda^{T-t-1}\underline{q})}{1-\lambda} > 0, \tag{33}$$

$$V_t^{bb}((\{\{\sigma^s\}_{t+1}^{T-1}, \{\sigma^p\}_T^\infty\})) - V_t^{ba}((\{\{\sigma^s\}_{t+1}^{T-1}, \{\sigma^p\}_T^\infty\})) = \frac{e((1-\lambda^{T-t-1})(1-\bar{q}) + \lambda^{T-t-1}(1-\underline{q}))}{1-\lambda} > 0, \tag{34}$$

(33) is increasing and (34) is decreasing in T. This implies that

$$\tau^a(q_t, \{\{\sigma^s\}_{t+1}^{T-1}, \{\sigma^p\}_T^\infty\}) - \tau^b(q_t, \{\{\sigma^s\}_{t+1}^{T-1}, \{\sigma^p\}_T^\infty\})$$

is increasing in T for all q. For the same initial condition $q_t > \bar{q}$, q_t is smaller the smaller is T for all $t > 0$ and q_T is smaller the larger is T. Notice that

$$\tau^a(q_t, \{\{\sigma^s\}_{t+1}^{T-1}, \{\sigma^p\}_T^\infty\}) - \tau^b(q_t, \{\{\sigma^s\}_{t+1}^{T-1}, \{\sigma^p\}_T^\infty\}) < \tau^a(q_t, \{\sigma^s\}_{t+1}^\infty) - \tau^b(q_t, \{\sigma^s\}_{t+1}^\infty)$$

for all T. There exist a finite $\hat{T}$ such $q_{\hat{T}-1} \geq \tilde{q}$, and $F(q_{\hat{T}-1}) \leq \tilde{q}$, since under $\{\sigma^s\}_{t+1}^\infty$, $\tilde{q}$ is reached in finite time.

>From 2.a) and 2.b) we conclude that $\{\{\sigma^s\}_{t+1}^{\hat{T}-1}, \{\sigma^p\}_{\hat{T}}^\infty\} = \sigma(q_{t+1})_{t\geq 0}$, and q_t converges to $\underline{q}$.

Case 3: $\underline{q} < \tilde{q} < \bar{q}$.

3.a) When $q_t < \tilde{q}$, $\{\sigma^p\}_{t+1}^\infty = \sigma(q_{t+1})_{t\geq 0}$ and q_t converges to $\underline{q}$.

3.b) When $q_t > \tilde{q}$, $\{\sigma^s\}_{t+1}^\infty = \sigma(q_{t+1})_{t\geq 0}$ and q_t converges to $\bar{q}$. □

Proof of Proposition 2

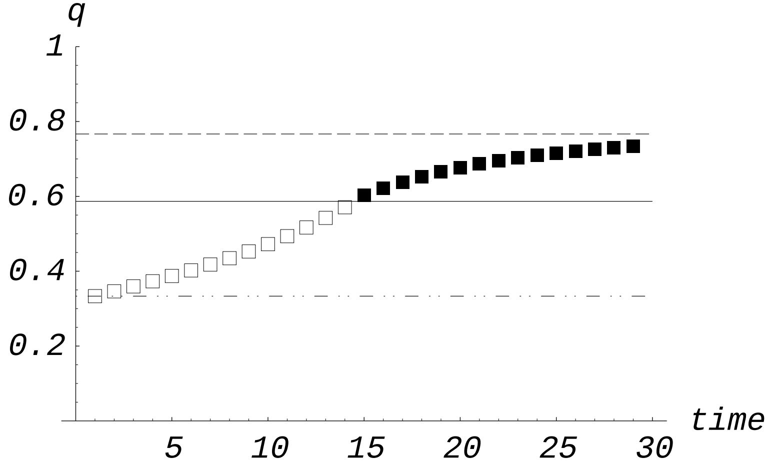
By lemma 2 $\tau^a(q_t, \{\{\sigma^p\}_{t+1}^{T-1}, \{\sigma^s\}_T^\infty\}) \gtreqless \tau^b(q_t, \{\{\sigma^p\}_{t+1}^{T-1}, \{\sigma^s\}_T^\infty\})$, when $q_t \lesseqgtr \underline{q} + \lambda^{T-t-1}(\bar{q} - \underline{q})$ Observe that

$$\underline{q} + \lambda^{T-t-1}(\bar{q} - \underline{q}) > \underline{q}$$

and that

$$\tau^a(q_t, \{\{\sigma^p\}_{t+1}^{T-1}, \{\sigma^s\}_T^{\infty}\}) - \tau^b(q_t, \{\{\sigma^p\}_{t+1}^{T-1}, \{\sigma^s\}_T^{\infty}\})$$

is decreasing in T for all q. For the same initial condition $q_0 = \underline{q} < \underline{q} + \lambda^{T-t-1}(\bar{q} - \underline{q})$, q_T is larger the larger is T (see 1.c in the proof to proposition 1). If we can find a T such that $q_{T-1} \leq \tilde{q} \leq q_T$, then $\sigma(q_{t+1})_{t\geq 0} = \{\sigma^p\}_{t+1}^{T-1}, \{\sigma^s\}_T^{\infty}\}$, and the proposed policy is credible. □



q
1
0.8
0.6
0.4
0.2
time
5
10
15
20
25
30