A Screening Role of Enforcement Institutions

Kenan Huremović*

June, 2016

Abstract

This paper examines scenarios where opportunistic behavior in contracts leads to the Prisoner's dilemma. In such instances, formal enforcement can encourage cooperation by raising the cost of defection. However, opting to cooperate even with weak enforcement signals unobservable characteristics, potentially enhancing desirability as a partner in the future. Consequently, the optimal welfare-maximizing enforcement quality may not always maximize cooperation. Depending on the distribution of unobservables, the impact of increased enforcement on cooperation can be influenced by reputation concerns. The payoffs of cooperators and defectors are non-monotonic in the quality of the enforcement institution. In cases where institutional quality is endogenously determined, the equilibrium level of enforcement tends to exceed the optimal level.

JEL: D02, D72, K42

Key Words: institutions, enforcement, matching, voting

CEAFE fields: Public economics, Law and Economics

*Aix-Marseille University (Aix-Marseille School of Economics), CNRS, & EHESS, 2 rue de la Charité, 13236 Marseille cedex 02, France. E-mail: kenan.huremovic@univ-amu.fr, Website: https://sites.google.com/site/kenanhuremovic

1 Introduction

Economic and social activities are governed by a set of formal and informal institutions. These institutions are important because markets and socio-economic activities in general cannot function well without them. As indicated in (Dixit, 2009) there are 3 main goals of governance institutions: (i) securing property rights (ii) enforcement of contracts (iii) resolving collective action problem. Here we focus on (ii), specifically on the problem of contractual opportunism. When the quality of institution that enforces contracts is low (in the sense that there is a low probability of being punished for breaking a contract), contractual opportunism creates a prisoner's dilemma problem and a society may end up in a state with very little or no cooperation at all. Some agents, however, might cooperate and not break a contract even when the quality of the enforcement institution is low due to their innate sense of righteousness or morals. Cooperating in this case signals that an agent has high morals or work ethics (from now on simply a type), which increases agent's reputation and makes her a better partner for future ventures. When the enforcement is extremely strong, cooperating will not signal anything about the type of agent. The same will be when the quality of enforcement institution is absent and costs of being defected on are high so that all agents will defect. In this case again no information about types of the players will be revealed. When the quality of legal enforcement is on some intermediate level such that some types choose to cooperate and others to defect, agent's action will reveal some information about her type.

This mechanism may have a significant effect in different contexts of human interaction. For example, consider a situation in which a population of firms interact in two different ways. The first type of interaction is the the simpler (routine) interaction which can be fully specified and thus monitored and enforced by a formal institution (for example, a delivery of specified products). The second one is a more complex and intangible interaction which cannot be monitored or enforced by a formal institution. Examples are R&D projects or joint ventures. In this type of activities an agent prefers having a partner with higher business ethics (better corporate culture, higher morals), as this will imply lower probability of opportunistic behaviour and therefore a higher expected payoff. Cooperating in the first type interaction will signal high business ethics when there is a possibility to defect and get away with it. The screening role of the enforcement institution discussed above can in this case facilitate positive assortative matching between firms with better business ethics in the second type interaction. When the expected pay-off from the second type activity exhibits complementarity, positive assortativeness will increase social welfare. A similar mechanism may be in play within the organization, concerning interaction in teams between employees. Less monitored and rigid work environment will facilitate more efficient team formation based on behaviour of employees in day to day operations; and thus more creativity, and output. Even in everyday interaction between people - not defecting when there is a chance to defect signals high moral values, making an agent a better partner in other socio-economic activities, such as marriage or different kinds of neighbourhood communal activities.

The interaction between formal incentives and signalling concerns has been discussed in the literature in different contexts and can be tracked back to (Titmuss, 1970), who argued that paying blood donors could reduce the supply as it makes donating blood 'less of a good deed' due to monetary incentives. This interaction between intrinsic (being a good, altruistic person) and extrinsic motivation is empirically and experimentally documented in different contexts. For example, (Gneezy and Rustichini, 2000) found that fining parents for picking up their children late from day-care centres resulted in more late arrivals, indicating that extrinsic motivation (fines) has crowded out intrinsic motivation (i.e. signalling being a good parent). In experiments, (Fehr et al., 2001) recorded that subjects provided less effort when the contract specified punishment for bad behaviour, compared to the case when it did not. The effect of extrinsic incentives on prosocial behaviour such as blood donation from theoretical point of view is discussed in (Benabou and Tirole, 2006). In the subsequent paper (Benabou and Tirole, 2011) the same authors in a similar framework analyse how laws and norms interact, specifically in the context of the optimal taxation problem. Even though the questions asked are different, the main idea in these papers is similar to the idea here: "The effect of the extrinsic motivation is substantially determined by the intrinsic motivation of an agent". In this context, we should also mention (Seabright, 2009), who considers a possibility of crowding out of intrinsic motivation in the model as in (Benabou and Tirole, 2006) with explicitly modelled signalling benefits, and (Levy and Razin, 2013) who studies how self and social signalling by being religious can increase the level of cooperation in a society.

In our model, signalling high type does not bring direct benefit, as in (Benabou and Tirole, 2006) for example, but has an instrumental role - it provides better match in the future. In this dimension, this paper has some similarities with the literature on costly signalling and matching (see for example (Hoppe et al., 2009)). However signalling here is not done by paying a cost directly, but by playing a certain strategy in a bilateral game. The game itself, together with the population of players, determines the cost (and the benefits) of signalling. Furthermore, which game will be played is determined by the quality of the enforcement institution. This, together with the distribution of types in the society, determines the cost of signalling.

Our paper is also related to the literature that studies the relationship between formal and informal enforcement institutions. The interaction between formal enforcement and reputation has been discussed in (Dixit, 2003). In his model, the gains from cooperation (honest trade) are greater the larger the distance between the pair of traders. On the the other hand, the frequencies of meetings and spread of information are locally biased. In this setting, reputation can sustain cooperation when the society is not too large. Otherwise, an external enforcement is needed to sustain cooperation. Other related papers that discuss interaction of formal and informal institution mostly focus on repeated interactions. The general conclusion is that better external enforcement crowds out the effect of informal institutions because it weakens reputation incentives (Kranton, 1996; Dhillon and Rigolini, 2011).

The paper is organized as follows. In Section 2 we present the basic model. The analysis of the equilibrium is conducted in Section 3, and in Section 4 we discuss the effects of the quality of the enforcement institution on the welfare. In Section 5 we analyze the quality of enforcement institutions when it is endogenously determined in majority voting elections. Section 6 discusses the effect of matching friction, and Section 7 concludes.

2 Setup

In the paper we are interested in bilateral interactions which, in the absence of enforcement institution, can be represented as a Prisonner's dilemma game (PD). The enforcement institution is modelled as a probability that a defector will be identified and fined. We refer to the magnitude of this probability as to the quality of (enforcement) institution. The quality of institution has a direct effect of increasing cooperation by making defection more costly and therefore less attractive.

We study a game with uncountable set of players. So, there is a population with a continuum of heterogeneous agents which differ in their cost of defecting. The cost of defecting is the type of an agent, and one can think of it as morals of an agent, work ethics, an ability to defect, a psychological cost of cheating or even to a some extent trust and trustworthiness[1].

The interaction happens across three periods. In period 1 players learn their types, and are randomly matched in pairs to play simultaneous move Bayesian game with payoffs as in Table 1. When both players cooperate, they receive payoff $\gamma > 0$. When player i defects and player j cooperates, the defector receives payoff $\kappa > \gamma$, but also experiences disutility $-\lambda_i$ (i.e. the psychological cost of immoral action). In addition to this, the defector is identified with probability $\theta \in [0, 1]$ and charged a fine normalized to 1. When this happens, the whole fine is transferred to the cooperator. Thus when player i defects and player j cooperates, the payoff of player i is $\theta(\kappa - 1 - \lambda_i) + (1 - \theta)(\kappa - \lambda_i) = \kappa - \theta - \lambda_i$. The cost of being defected on when cooperating is $c > 0$. In this case player j suffers a loss c and receives transfer from the defector of size 1 with probability θ. We refer to θ as to the quality of the enforcement institution.

	C	D
C	γ, γ	$\theta - c, \kappa - \theta - \lambda_j$
D	$\kappa - \theta - \lambda_i, \theta - c$	$-\lambda_i, -\lambda_j$

Table 1: Two players game

Agents don't know the type of the agent they are matched with, but they do know the distribution of types and own type. So type of an agent is private information. In period 2 payoffs from period 1 are realized and actions taken by the every player in the first period are observed by everyone[2]. In period 3, depending on the action in the game from period 1, agents are assortatively matched with other players who acted in the same way (cooperators with cooperators and defectors with defectors[3]). Payoff from the third period is assumed

[1]In this context an agent is said to be more trustworthy if the probability that she will defect conditional on opponent cooperating is low. An agent will be more trusting if it is more likely that she will cooperate

[2]Check Section 6 for the discussion on the observability assumption

[3]We shall allow for friction with this respect in section 6

to exhibit complementarity in types in a form standard in the matching literature. When players i and j are matched, the benefit to both of them is $\mu\lambda_i\lambda_j$ where $\mu > 0$ is the parameter measuring the importance of the matching stage. The discount factor is built in μ. There are no strategic decisions in the third period, and one can think of it as a reduced form representation of a potentially long term bilateral project that is intangible or too complex to be enforced by a formal institution (such as an entrepreneurial or R&D project). This formulation captures the idea that expected benefit from the joint project is higher if a partner has higher cost of defection (higher morals, work ethics).

2.1 Parameters

Throughout the paper we shall maintain some assumptions on the parameter space which we discuss in this section. First, for simplicity, we assume that the support for type distribution is $\left[\underline{\lambda}, \overline{\lambda}\right] = [0, 1]$. We will also assume that the cooperation is efficient from the single shot game perspective in the first period, even for the worst type($\lambda = \underline{\lambda}$) - that is that $2\gamma > \kappa - c$. We shall also postulate that when $\theta = 0$, meaning that practically there is no legal enforcement, the defection is the optimal strategy for every player. This means that when $\theta = 0$ the game is PD. In terms of parameters this means that: $\kappa - 1 > \gamma \wedge -c < -1 \Rightarrow \kappa > \gamma + 1 \wedge c > 1$

To be more clear let us state the assumptions explicitly:

Assumption 1. *We assume the following restrictions on parameters*

- $2\gamma > \kappa - c$ - *Cooperation is efficient*
- *The game is PD when* $\theta = 0$
  - *(i)* $\gamma < \kappa - 1$
  - *(ii)* $c > 1$

The assumption 1 gives $\kappa - c < 2\gamma < 2\kappa - 2 \Rightarrow \kappa > 2 - c$

For the sake of the discussion in sections below, let us define the concept of strategic substitutes and strategic complements formally, following (Bulow et al., 1985).

Definition 1. *A bilateral game is said to be a game of strategic complements (substitutes) if for every two players i and j and their strategies x_i and x_j:* $\frac{\partial^2 \pi_i}{\partial x_i \partial x_j} > 0$ *(< 0)*

From Definition 1 it follows that in a game of strategic complements (substitutes) best response functions are upward (downward) sloping. In the model considered in our case, strategic complements (substitutes) will imply that player is more (less) prone to cooperate when the probability of being matched with cooperator is higher.

3 Equilibrium

Note that in the third period (matching stage) there are no strategic decisions. Thus, the game here is basically a simultaneous move game with incomplete information, and therefore we shall employ the concept of Bayesian-Nash equilibrium (BNE). Let us consider a match between arbitrary agents i and j, and let p_j denote the probability that player j will cooperate. We can write the expected payoff of cooperation of player i as:

$$\pi(\lambda_i, C) = \gamma p_j + (\theta - c)(1 - p_j) + \mu\lambda_i E(\lambda|C)$$

The expected payoff of defecting is:

$$\pi(\lambda_i, D) = (\kappa - \theta - \lambda_i)p_j + (-\lambda_i)(1 - p_j) + \mu\lambda_i E(\lambda|D)$$

where $E(\lambda|C)$ is the expected type of cooperators, and $E(\lambda|D)$ is the expected type of defectors. Thus the net benefit of cooperating is:

$$\Pi(\lambda_i, \lambda_j) = \pi(\lambda_i, C) - \pi(\lambda_i, D) = (\gamma - \kappa + c)p_j + \lambda_i - c + \mu\lambda_i \left(E(\lambda|C) - E(\lambda|D)\right) \quad (1)$$

As $\lambda_i \in [0, 1]$, the strategy of player i is a function $\sigma_i : [0, 1] \to \{C, D\}$. We say that player i uses cutoff strategy, if there exists some $x \in [0, 1]$ such that:

$$\sigma_i(\lambda_i) = \begin{cases} \text{C} & \text{if } \lambda_i \geq x \\ \text{D} & \text{otherwise} \end{cases}$$

A cutoff equilibrium is BNE in cutoff strategies.

Suppose that player j uses cut-off strategy with cutoff y ($\sigma_j(\lambda_j) = y$). Than the probability of cooperation of player j is $1 - F(y)$ where F is cdf of the type distribution. We write:

$$\Pi(\lambda_i, y) = (\gamma - \kappa + c)(1 - F(y)) + \lambda_i - c + \mu\lambda_i \left(E(\lambda|C) - E(\lambda|D)\right)$$

For cutoff strategy to be a best response it must be that $\Pi(\lambda_i, y)$ is increasing in the first argument, that is, player i should be more prone to choose C over D the more ethical she is.

We have that $\frac{\partial \Pi(\lambda_i, y)}{\partial \lambda_i} = 1 + \mu \left(E(\lambda|C) - E(\lambda|D)\right)$ which will always be greater than 0 given that all players use the cutoff strategy. Given the positive sign of the partial derivative, we can write the best response of player i to cutoff strategy y of player j as:

$$\sigma_i(y) = \begin{cases} 0 \text{ if} & \Pi(0, y) \geq 0 \\ 1 \text{ if} & \Pi(1, y) \leq 0 \\ x \in (0, 1) & \text{otherwise} \end{cases}$$

where x is the unique solution of equation $\Pi(\lambda, y) = 0$ (so when $\lambda_i > x$ player i plays C and otherwise plays D)

As players are ex ante symmetric, we shall look for the symmetric equilibrium in cutoff strategies (in which all players choose the same cutoff).

The following proposition gives a result regarding existence of the cuotoff equilibria:

Proposition 1. *There exist a cutoff equilibrium*

Proof. See Appendix A □

Let us examine the effect of the quality of enforcement institution (θ), and reputation $\mu\Delta(\lambda)$ on the level of cooperation in the equilibrium. Taking the derivative of(17) with respect to λ we get:

$$\frac{\partial \lambda^*(\theta)}{\partial \theta} = -\frac{1}{1 + \mu\lambda\Delta'(\lambda) - (\gamma - \kappa + c)F'(\lambda)} \tag{2}$$

which for the region of parameters when λ^* is unique (Proposition 3) is always positive. Similarly we get that $\frac{\partial \lambda^*}{\partial \mu} > 0$. Thus the increasing quality of the enforcement institution and relative importance of the reputation benefits will increase the cooperation in the first period. However, increasing θ, may actually decrease the reputation benefits for cooperating vs. defecting. Before discussing this, let us first state the result relating density function f with function Δ. Let f be a density function $f : \left[\underline{\lambda}, \overline{\lambda}\right] \to \mathbb{R}$. Let us define function $\Delta : \left[\underline{\lambda}, \overline{\lambda}\right] \to \mathbb{R}$ with

$$\Delta(x) = E(\lambda|\lambda > x) - E(\lambda|\lambda < x)$$

where expectation is according to density function f. Then the following result holds (Jewitt, 2004):

Theorem 1 (Jewitt). *If f is everywhere decreasing (increasing) then Δ is everywhere increasing (decreasing). When f is unimodal, Δ is quasiconvex. If f has a unique interior maximum then Δ has an unique interior minimum*

Proof. Omitted □

Following Theorem 1 we can state the following result:

Proposition 2. *When distribution of types has increasing (decreasing) density function $f : [\underline{\lambda}, \overline{\lambda}] \to \mathbb{R}$, the reputation benefit of cooperating versus defecting in the interior equilibrium will increase (decrease) with the quality of enforcement θ. When f reaches its maximum in the interior of $[\underline{\lambda}, \overline{\lambda}]$, the reputation benefit of cooperating versus defecting will reach its minimum in the interior $[\underline{\lambda}, \overline{\lambda}]$*

Proof. Follows directly from Theorem 1. □

For instance, a decreasing density function describes a situation in which there are more 'bad types' (agents with low cost of defecting) relative to the number of 'good types' (i.e. in the case of decreasing density function) in the society. The effect of improving the quality of enforcement institution on the level of cooperation will then be partially crowded out, due to the decrease of the reputation incentive to cooperate vs. defect. Increasing θ will decrease equilibrium threshold $\lambda^*(\theta)$ and increase the mass of agents that cooperate. As f is decreasing, this means that the expected type of cooperator will decrease more than the expected type of defector, causing decrease in $\Delta(\lambda^*)$. This will partially crowd out the effect of increase of θ on the level of cooperation. The strength of the crowding out effect is captured by the term $1 + \mu\Delta'(\lambda^*)$ in the equation (2). When the density function is increasing, the reputation effect will reinforce the direct effect of θ. This is in essence the same mechanism as in (Benabou and Tirole, 2006) applied to the situation modelled in this paper.

Another, rather intuitive, insight the from equation 2 is that the effect of increase θ on the level of cooperation will be smaller in the case of strategic substitutes ($\gamma - \kappa + c < 0$), than in the case of strategic complements ($\gamma - \kappa + c > 0$).

In what follows we shall focus on the parameter space in which the cutoff equilibrium is unique. The following proposition gives a result regarding that:

Proposition 3. *The cutoff equilibrium will be unique when:*

$$\mu\Delta(y) + 1 > |\mu\sigma(y)\Delta'(y) - (c + \gamma - \kappa)F'(y)|$$

Proof. See Appendix A □

3.1 Analysis of the Equilibrium

In this section we shall assume that the distribution of types is uniform on the segment $[0, 1]$. The indifference condition can be be written as:

$$\Pi(\lambda, \lambda) = (c + \gamma - \kappa)(1 - \lambda) - c + \theta + \frac{\lambda}{2} + \lambda = 0 \tag{3}$$

which gives the cutoff point:

$$\lambda^* = \frac{\kappa - \gamma - \theta}{\kappa - \gamma - c + 1 + \frac{\mu}{2}} \tag{4}$$

Let us first state the following corollary of Proposition 3, defining a region of parameters for which the threshold equilibrium is unique.

Corollary 1. *When types are distributed uniformly on the segment* $[0, 1]$ *the equilibrium will be unique when*

$$\frac{1}{2}(\mu + 2) > |\gamma - \kappa + c|$$

Proof. When types are uniformly distributed on $[0, 1]$ we get:

$$\sigma_i'(y) = 2\frac{\gamma - \kappa + c}{\mu + 2} \tag{5}$$

and the claim directly follows from Proposition 3 □

First note that the threshold defined with equation (4) is a linear function of θ. By Proposition 3 the denominator of (4) will be positive, and therefore the threshold will always decrease with θ (mass of cooperators will increase). The threshold will also decrease with μ as: $\frac{\partial \lambda^*}{\partial \mu} = -\frac{2(\kappa - \gamma - \theta)}{(-2c - 2\gamma + 2\kappa + \mu + 2)^2} > 0$ since by the Assumption 1 $\kappa - \gamma - \theta > \kappa - \gamma - 1 > 0$.

The threshold will be interior when $\frac{\mu}{2} + 1 - c > -\theta \Leftrightarrow c < 1 + \theta + \frac{\mu}{2}$, thus when the cost when defected on is not too high, or when the long term benefits, captured with μ are relatively high. Note that due to Assumption 1 we have that $\kappa - \gamma > \theta$ as $\theta \in [0, 1]$ so the expression in the numerator of (4) is always positive. Thus we have showed that the following lemma holds.

Lemma 1. *In the case of uniform distribution of types, the threshold will be interior when*

$$c < 1 + \theta + \frac{\mu}{2} \tag{6}$$

In what follows we shall focus on the cases when the threshold is interior.

Let us check now how the payoffs depend on the level of enforcement. For type λ (and interior threshold) the payoff of the cooperator in equilibrium is given with:

$$\pi(\lambda, C) = \frac{2\lambda\mu(c + 2\gamma + \theta - 2\kappa) + 4(c - \theta)(\kappa - \theta) - 2\mu(\gamma + \lambda) - 4\gamma - \lambda\mu^2}{4(c + \gamma - \kappa - 1) - 2\mu} \tag{7}$$

and the payoff of defector with type λ at interior threshold is:

$$\pi(\lambda, D) = \frac{-2\lambda(c + \gamma - \kappa - 1) + 2(c - \theta - 1)(\kappa - \theta) + \mu(\lambda(\gamma + \theta + 1) + \theta - \kappa(\lambda + 1))}{2(c + \gamma - \kappa - 1) - \mu} \tag{8}$$

There are a couple of interesting things worth noting here. First let us look at the payoff of cooperators. Taking derivative with respect to θ we get:

$$\frac{\partial\pi(\lambda, C; \theta)}{\partial\theta} = \frac{2(c - 2\theta + \kappa) - \lambda\mu}{-2c - 2\gamma + 2\kappa + \mu + 2} \tag{9}$$

We are interested in the behaviour at the unique cutoff equilibrium, thus due to Proposition 3, we have that denominator of (9) will always be positive. The sign of the numerator $2(c - 2\theta + \kappa) - \lambda\mu$ can be either positive or negative. However, if this expression is negative for some $\lambda = \tilde{\lambda}$ it is negative for all $\lambda > \tilde{\lambda}$. That is, if an increase in the quality of institution decreases payoff of player with type $\tilde{\lambda}$ it will decrease the payoff for every other player that has higher type (higher morals). The expression (9) will be negative when $\lambda > \frac{2c-4\theta+2\kappa}{\mu}$. Also, from Assumption 1 it follows that $c - 2\theta + \kappa > 0$[4]. So for low values of μ this expression will be positive, and it becomes negative when μ is higher. The intuition is that if μ is high, the benefit from the matching stage is high and it becomes more important to be matched with higher type. When θ increases then by (4) the threshold will decrease, increasing the number of cooperators, and thus the payoff from the first period. However, this will result in a decrease of the average cooperator type, and in turn decrease the expected payoff of cooperator from the matching stage. Which effect will dominate depends on the size of μ. We can now state the following proposition:

Proposition 4. *Increase in the quality of enforcement institution will decrease the benefit of cooperator with type* $\lambda > \frac{2c-4\theta+2\kappa}{\mu}$. *This will be the case when reputation benefits are high enough, so* $\mu \geq \frac{2(c+\gamma-\kappa-1)(c-2\theta+\kappa)}{c+\gamma-\theta}$

Proof. See Appendix A □

Let us look at the payoff of a defector. Partial derivative with respect to θ is

$$\frac{\partial\pi(D, \lambda, \theta)}{\partial\theta} = \frac{2\kappa + 2c - 4\theta - 2 - \mu(1 + \lambda)}{-2c - 2\gamma + 2\kappa + \mu + 2} \tag{10}$$

As was the case with cooperator, the partial derivative is monotone in λ. When both θ and μ are low, the sign of (10) might be positive. This is because an increase in θ will increase the

[4] As $c > 1 \wedge \kappa > \gamma + 1 > 1 \wedge \theta \in [0, 1]$

mass of cooperators, by decreasing the threshold. This will increase the expected payoff of defectors from the first period, as it increases the probability that they will be matched with a player who cooperates in the equilibrium. Increase in θ will make the average defector type lower, hence lowering the expected payoff from the second stage for defectors. This effect will not be strong if μ is low enough, so we can state:

Proposition 5. *The payoff of defector of type λ can increase with the quality of enforcement institution. This will be the case when type λ is defector and $\lambda < \frac{2\kappa+2c-4\theta-2-\mu}{\mu}$, which is satisfied when reputation benefit is not too high*

Proof. See Appendix A □

What will be the effect of the increase of quality of enforcement institution is determined by the value of the matching in the third stage. If the matching is relatively unimportant, then the main effect of increasing θ for the payoff of cooperators will be through the increase the number of cooperators (thus the probability of meeting a cooperator) in the first stage, which will be positive. As for defectors, increasing the number of cooperators will increase the expected payoff, as it increases the probability of meeting the cooperator. On the other hand, increase in θ will increase the expected fine when defecting. The expected payoff from the third period for both cooperator and defector will be smaller when threshold decreases. This is because increase in the threshold will decrease the expected type of cooperators and the expected type of defectors. If μ is high, this effect will dominate and payoff of every player will decrease with θ (given that threshold is interior). When μ is small, an increase in θ will increase the payoff of cooperators and the effect on the payoff of defector will depend on the behaviour of threshold $\lambda^*(\theta)$. The effect of changing the quality of the enforcement institution thus can have different effects on the payoffs of agents in the society. In the next section we examine what will be the effect on the total welfare.

4 Welfare

From the first stage perspective it is optimal to maximize the cooperation, as it is the efficient outcome of the first period game (first stage). However, the matching in the third period is done based on actions taken in the first stage. As the payoff from the matching in period 3 exhibits complementaries, the efficient matching is the perfect positive assortative matching. However, the type of an agent is a private information, and only thing that is observable is the action in the first stage game. Thus the matching will be coarse, with only two classes of agents - cooperators and defectors from the first stage. Two classes seems to be too coarse to get a significant gain in welfare, compared to the purely random matching. However, it has been shown in McAfee (2002) that with two class matching with uniform distribution 75% of the value of using infinitely many classes can be obtained. When both are one-tailed exponential, with two classes 74.62% of possible gains can be obtained. When both are normally distributed, two classes result in about 63% of the possible gains.

In our case, the matching with two classes and a cutoff point of λ^* from the same population means associating values below λ^* with values below, and values above λ^* with values above λ^*. In this case, the social value of the matching is:

$$W_M = F(\lambda^*) \left(\int_{\underline{\lambda}}^{\lambda^*} \lambda \frac{f(\lambda)}{F(\lambda^*)} \, \mathrm{d}\lambda \right)^2 + (1 - F(\lambda^*)) \left(\int_{\lambda^*}^{\bar{\lambda}} \lambda \frac{f(\lambda)}{1 - F(\lambda^*)} \, \mathrm{d}\lambda \right)^2 \tag{11}$$

Players with type lower than λ^* (defectors) are matched with the players that have type lower than λ^*, and cooperators are matched with cooperators. Thus, the average match value is the probability that a player is defector times the expected value of the defector match, plus the analogue term conditional on both being cooperators (having type higher than λ^*). Here λ^* is defined with (4). The matching here is perfect in the sense that cooperators are matched with cooperators and defectors are matched with defectors.

To find the optimal threshold λ^* (optimal θ that induces this λ^*) social planner is facing a trade-off. More cooperation will increase the total payoff from the first stage, but it might not reveal enough information and thus make matching less efficient. Maximizing social gains from matching requires a loss in the first stage, due to the fact that then there must be some defectors in order to better screen out the high types. When μ is small compared to the benefits from cooperation in the first stage, than the optimal λ^* will be the one that maximizes cooperation. When μ is larger, then matching payoff contributes relatively more to the welfare, and thus it will be optimal to have λ^* interior. The following proposition characterizes the optimal value of θ^*

Proposition 6. *For $2\gamma + c - \kappa \geq \frac{\mu}{8}$, $\theta_u^* = 1$ maximizes welfare. Otherwise the welfare will be maximized with*

$$\theta = \theta_u^* = \frac{3(4\gamma + 1)(c + \gamma - \kappa)}{-8c - 8\gamma + 8\kappa + \mu + 2} + \frac{1}{2}(c + 3\gamma + \kappa) - \frac{\mu}{4}$$

which defines the threshold:

$$\lambda^* = \frac{8c + 16\gamma - 8\kappa - \mu}{2(8c + 8\gamma - 8\kappa - \mu - 2)}$$

Proof. See Appendix A □

When the social benefit from complementarity in the third period is small compared to the benefits from the first stage i.e. the society gain from the cooperation, then it is optimal to minimize mass of defectors in the society, setting $\theta^* = 1$. When μ is large, then it will be optimal to have some 'extra' defectors in the society, in order to more accurately screen out the good types for the matching, and thus $0 < \theta_u^* < 1$

5 Endogenous Quality of Enforcement

Having in mind the social game context discussed in the introduction, it is natural to ask what level of θ will emerge endogenously from a process of decentralized social decision making. The question is even more interesting having in mind that the results from Section

3.1 state that the payoff of defector and cooperator can be both increasing and decreasing in θ.

When deciding on his preferred level of the enforcement, an agent faces a tradeoff. Conditional on cooperating, a player would like θ to be higher to protect him from being defected on in period 1. On the other hand, he prefers a higher average type of cooperators in the equilibrium as then the reputation benefit will be higher, which happens when θ is lower. Conditional on defecting, an agent prefers lower θ as this implies lower probability of being identified and fined as a defector. On the other hand, he would like θ to be high enough so there are still some cooperators in the society that he can exploit in the first period. Furthermore, as higher θ implies lower average type of defectors, he would prefer θ to be lower, as this will increase his benefits from the third period. Thus, which θ will emerge in a decentralized social decision making process is determined by a complex interaction between reputation concerns and the direct effect of θ in changing rules of interaction in the first period.

To be more concrete, in this paper we shall focus on a simple voting model in which agents vote sincerely on the value of θ, and study how does the equilibrium θ depends on the primitives of the model. Individual preferences over θ are single peaked, and as the voting is done over one-dimensional issue (value of θ) we can apply the median voter theorem. The agents will vote differently conditional on whether they cooperate or defect in the equilibrium.

Type λ_i will prefer different values of θ conditional on cooperating or defecting, and let us denote these two values with θ_i^C and θ_i^D respectively. If payoff of player i is larger when defecting and $\theta = \theta_i^D$ than when cooperating and $\theta = \theta_i^C$, then she will opt for θ_i^D.

The following propostion gives the value of θ in the voting equilibrium:

Proposition 7. *The majority rule voting system will chose level of enforcement* $\theta_v^* = \frac{c+\kappa}{2} - \frac{\mu}{8}$

Proof. See Appendix A □

Let us compare θ chosen by the median voter stated in Proposition 7 (θ_v^*), with θ that maximizes the social welfare, defined in Proposition 6 (θ_u^*). We have that:

$$\begin{aligned}
\theta_v^* - \theta_u^* &= \frac{c+\kappa}{2} - \frac{\mu}{8} - \left(\frac{3(4\gamma+1)(c+\gamma-\kappa)}{-8c-8\gamma+8\kappa+\mu+2} + \frac{1}{2}(c+3\gamma+\kappa) - \frac{\mu}{4} \right) \\
&= -\frac{3(4\gamma+1)(c+\gamma-\kappa)}{-8c-8\gamma+8\kappa+\mu+2} - \frac{3\gamma}{2} + \frac{\mu}{8} \\
&= \frac{-\mu^2 + (8c+20\gamma-8\kappa-2)\mu + 24c + 48\gamma - 24\kappa}{8(8c+8\gamma-8\kappa-\mu-2)} \qquad (12)
\end{aligned}$$

The derivative of the expression (12) with respect to μ is given with the expression:

$$\frac{3(4\gamma+1)(c+\gamma-\kappa)}{(-8c-8\gamma+8\kappa+\mu+2)^2} + \frac{1}{8} \qquad (13)$$

which is always positive in the case of strategic complements $\gamma - \kappa + c > 0$. Furthermore, due to the uniqueness condition in Proposition 3 we have that at the left side of the parameter space with respect to μ expression (12) will have value: $\frac{1}{4}(c + 3\gamma - \kappa + 1)$. Thus, (12) is increasing in μ and takes positive value at the lowest feasible value of μ. Thus (12) will always be positive. Note that in the strategic complements case γ is always relatively low comparing to μ.

In the case of strategic substitutes $\gamma - \kappa + c < 0$ it might as well be that $\theta_v^* < \theta_u^*$, when μ is small relative to γ. When γ is high comparing to μ it becomes more important to have larger mass of cooperators in the society and the screening role of θ becomes less important. This means that θ_u^* will be higher. However, the quality of enforcement institution θ_v^* preferred by the median voter does not depend on γ. When γ is high enough, and μ low, then provided that we are in the case of strategic substitutes, $\theta_u^* \geq \theta_v^*$. As was the case with strategic complements, in the case of strategic substitutes the difference $\theta_v^* - \theta_u^*$ is increasing in μ, as (13)will be always larger than zero. Even though the first part of the expression (13) is negative, the quadratic equation

$$\frac{1}{8} - \frac{3(4\gamma + 1)(c + \gamma - \kappa)}{(-8c - 8\gamma + 8\kappa + \mu + 2)^2} = 0$$

doesn't have real zeros, and will always have the positive value (in the considered parameter space). As μ increases, the matching stage becomes more important, decreasing the optimal θ_u^* faster than it decreases the optimal θ_v^*, eventually making $\theta_u^* \leq \theta_v^*$

6 Matching Frictions

So far we have assumed that there is perfect matching in the third period in a sense that cooperators will always be matched with cooperators and defectors will always be matched with defectors. In this section we discuss the case when this isn't necessarily true. We assume that there is a non-negative probability that a cooperator is matched with defector in the third period. Let $P(C|\tau)$ denote the probability of a player being matched with cooperator in the third period given that his action in the first period is $\tau \in \{C, D\}$. In this case the following balancing condition must hold:

$$(1 - F(\lambda^*))(1 - P(C|C)) = F(\lambda^*)P(C|D) \tag{14}$$

The condition (14) states just that the mass of defectors that are matched with cooperators is equal to the mass of cooperators matched with defectors. As all players are matched in the third period, it holds that $P(C|C) + P(D|C) = 1$. Let us denote with $\alpha(\lambda^*) = P(C|C) - P(C|D)$. In the matching literature $\alpha(\lambda^*)$ is referred to as 'index of assortativity' (Bergstrom, 2003). Furthermore, let $\zeta = P(C|C)$ and $\xi = P(C|D)$, then $\alpha = \zeta - \xi$. Using this notation we can write(14) as:

$$(1 - F(\lambda^*))(1 - \zeta)) = F(\lambda^*)\xi \Rightarrow$$

$$\xi = \frac{(1 - F(\lambda^*))(1 - \zeta))}{F(\lambda^*)} \Rightarrow$$

$$\alpha = \zeta - \frac{(1 - F(\lambda^*))(1 - \zeta))}{F(\lambda^*)} = \frac{\zeta + F(\lambda^*) - 1}{F(\lambda^*)}$$

Then the expected reputation benefit of cooperator and defector of type λ_i denoted with $R(\tau, \lambda)$ $\tau \in \{C, D\}$: can be written as:

$$\begin{aligned}
R(C, \lambda_i) &= \lambda_i \left(\zeta E(\lambda|C) + (1-\zeta)E(\lambda|D)\right) \\
R(D, \lambda_i) &= \lambda_i \left(\xi E(\lambda|C) + (1-\xi)E(\lambda|D)\right) \\
&= \lambda_i \frac{(1-F(\lambda^*))(1-\zeta))}{F(\lambda^*)} E(\lambda|C) + \left(1 - \frac{(1-F(\lambda^*))(1-\zeta))}{F(\lambda^*)}\right) E(\lambda|D)
\end{aligned}$$

And reputational benefit from cooperating vs. defecting is then:

$$\begin{aligned}
R(C, \lambda_i) - R(D, \lambda_i) &= \lambda_i \left((\zeta - \xi)E(\lambda|C) + (1 - \zeta - 1 + \xi)E(\lambda|D)\right) \\
&= \lambda_i(\zeta - \xi) \left(E(\lambda|C) - E(\lambda|D)\right) \\
&= \lambda_i \alpha \left(E(\lambda|C) - E(\lambda|D)\right)
\end{aligned}$$

where α is the index of assortativity. We shall be interested in the case of positive assortative matching in the third period, thus for the values of $\alpha \in [0, 1]$. Note that then the analysis is parallel with what we had before, with μ being α scaled by some positive factor.

In this context it might be more natural to conduct the analysis with respect to ζ - the probability that an agent will be matched with cooperator given that he cooperates. ζ can be interpreted as a speed of information transmission in the social network, or simply as the quality of informal enforcement institution (reputation). Keeping ζ fixed, increase of θ, will decrease λ^*, $F(\lambda^*)$, and this will make assortativity index higher for a fixed ζ, increasing further on the level of cooperation. This effect captures the complementarity between formal enforcement and reputation throughout a different channel than it has been discussed in context of function Δ.

For example, in the case of uniform distribution (3) becomes:

$$\Pi(\lambda, \lambda) = (c + \gamma - \kappa)(1 - \lambda) - c + \theta + \mu \frac{\zeta + \lambda - 1}{\lambda} \frac{\lambda}{2} + \lambda = 0 \Leftrightarrow \tag{15}$$

$$\Pi(\lambda, \lambda) = (c + \gamma - \kappa)(1 - \lambda) - c + \theta + \frac{\mu}{2} (\zeta + \lambda - 1) + \lambda = 0 \tag{16}$$

and solving we get:

$$\lambda^* = \frac{-2\gamma - \zeta\mu - 2\theta + 2\kappa + \mu}{-2c - 2\gamma + 2\kappa + \mu + 2}$$

from where it is visible that λ^* will decrease with ζ. The further analysis parallel to what has been done in the frictionless matching case in the paper can be conducted in analogous way, and the qualitative nature of the results will not change.

7 Conclusion and Extensions

A higher level of enforcement and monitoring will make agents behave according to a prescribed set of rules. However, there are a lot of situations in which the tasks are too complex or intangible to be enforced in such a way. In these situations it is essential to find a partner who will exert high effort, even when not being subject to formal enforcement. When the returns from more complex interactions are higher than from the interactions that can be monitored, it is socially optimal to have lower level of enforcement such that some agents will choose to break the rules, and by doing that signal that they have bad work ethics. This will screen out agents who follow the rules only because of the high enforcement (disutility of being punished) from agents who choose to take a prescribed action as they feel intrinsic disutility from defecting. This will facilitate more efficient matching in tasks that cannot be enforced by a set of formal rules. The model captures this screening role of enforcement institution - emphasising the (nonmonotonic) effect of the strength of the enforcement on welfare and payoffs of agents with high and low work ethics. The set of rules in the society usually arises endogenously through some kind of social decision process. How the equilibrium institutions compare to the optimal depends on the importance of the screening role of the enforcement institution and whether the interaction in the monitored activities has the nature of strategic complements or strategic substitutes.

The formal (θ) and informal (reputation function Δ) enforcement institution interact in determining the level of cooperation and social welfare. By behaving according to the rules an agent earns a reputation to be of high work ethics. The benefit from the good reputation will be higher if a smaller mass of other agents behave the same way - and thus depends and the level of formal enforcement. The matching distortions will decrease the reputation benefit.

The model provides a good framework to study interaction between formal institution (enforcement) and informal institution (reputation) in enhancing cooperation in the society and the effects on the welfare. In relation to this, we find that the effect of the formal enforcement can be both reinforced or diminished by the reputation concerns. When distribution of types is non-decreasing, increase in the quality of institution will increase the reputation benefits, and thus the effect of reputation and institutional quality will go in the same direction, increasing the level of cooperation. When the density function of type distribution is decreasing, then increase of the quality of institution will increase the level of cooperation, but the effect will be partially crowded out by reputation - as net reputation benefit of cooperation will decrease. When the distribution of types is not monotonic, the direction of the reputation effect will depend on the current level of θ. Lower matching frictions, as jet another dimension of reputation as informal institution, imply higher level of cooperation. A higher quality of enforcement institution can increase the quality of matching, increasing the level of cooperation in the society even further.

There are several direction to extend the current model, and some of them are already work in progress. First we have focused on the reputation effect on the intensive margin. The agents do not make decisions to participate in the long term (complex) projects, nor which action to take. The matching stage is just a reduced form representation of this type of interaction, chosen to capture the basic idea that interacting with an agent of higher work ethics will bring higher benefit. Extending the model by allowing agents to choose to engage in non monitored activities that can yield higher returns (i.e. entrepreneurial project), or to interact in enforceable activities, would allow us to examine the information role of the enforcement institution on the level of entrepreneurship in a society on the extensive margin.

This can, for example, have an important implications in studying the interaction between quality of enforcement institutions and the level of entrepreneurship in the society.

Finally, the level of enforcement affects the utility of cooperators and defectors in a non monotonic way. For some levels of θ material payoff of defectors is higher than payoff of cooperators and vice versa. In a simple indirect evolutionary process, akin to the one in (Ok and Vega-Redondo, 2001), the level quality of enforcement institution will determine the distribution of types in the society. On the other hand, agents choose the level of enforcement in elections as described in the paper, and the equilibrium enforcement level will depend on the distribution of types. Examining the co-evolution of institution and the type distribution in the society seems like interesting avenue for building on this model.

8 Appendix A: Proofs

Proof of Propostion 1: $\Pi_i(x, y)$ is increasing in x for all $y \in [0, 1]$ and $\Pi_i(\lambda, \sigma(\lambda))$ is continuous in λ. When $\Pi_i(1, 1) \leq 0$ then 1 is the equilibrium cutoff. If $\Pi_i(0, \sigma(0)) \geq 0$, 0 is an equilibrium. Otherwise the equilibrium is defined as a solution of equation $\Pi_i(\lambda, \lambda) = 0$, that is given with:

$$(c + \gamma - \kappa)(1 - F(\lambda)) - c + \theta + \lambda + \lambda\mu\Delta(\lambda) \tag{17}$$

□

Proof of Proposition 3: To show uniqueness it is sufficient to show that the slope of the best response is in absolute value smaller than 1. It is known that this guarantees that the two best response curves can intersect only once, see (Vives, 2000). The slope of the best response function $\sigma(y)$ we obtain by equating (17) to zero and taking implicit derivation. That is we have:

$$\sigma_i'(y) = -\frac{\Pi_2(\sigma_i(y), y)}{\Pi_1(\sigma_i(y), y)} = -\frac{\mu\sigma(y)\Delta'(y) - (c + \gamma - \kappa)F'(y)}{\mu\Delta(y) + 1} \tag{18}$$

□

Proof of Proposition 4: From (9) equilibrium payoff cooperator will decrease with θ when $\lambda \geq \tilde{\lambda} = \frac{2c-4\theta+2\kappa}{\mu}$. Type $\tilde{\lambda}$ will be cooperator when $\tilde{\lambda} \geq \lambda^*$. So the equilibrium payoff of cooperator of type $\tilde{\lambda}$ will decrease with θ when $\frac{2c-4\theta+2\kappa}{\mu} \geq \frac{\kappa-\gamma-\theta}{\kappa-\gamma-c+1+\frac{\mu}{2}}$ which will be the case when $\mu \geq \frac{2(c+\gamma-\kappa-1)(c-2\theta+\kappa)}{c+\gamma-\theta}$. When payoff of $\tilde{\lambda}$ decreases with θ, so does the payoff of type $\lambda > \tilde{\lambda}$. □

Proof of Proposition 5: From (10) equilibrium payoff of defector of type λ will increase with θ when $\lambda \leq \tilde{\lambda} = \frac{2\kappa+2c-4\theta-2-\mu}{\mu}$. Type $\tilde{\lambda}$ will be defector when $\tilde{\lambda} < \lambda^*$. So the type $\tilde{\lambda}$ will be defector and have payoff which will be increasing function of θ when $\frac{2\kappa+2c-4\theta-2-\mu}{\mu} < \lambda^*$ When payoff of type $\tilde{\lambda}$ is increasing with θ, so is the payoff of every type $\lambda < \tilde{\lambda}$ □

Proof of Proposition 6: In our case, the distribution is by assumption uniform on $[0, 1]$ which gives the following expression for social gain from matching in the second stage, where λ^* is the threshold dividing population into two classes.

$$W_M = \mu\left(\lambda^*\left(\int_0^{\lambda^*}\lambda\frac{1}{\lambda^*}\,\mathrm{d}\lambda\right)^2 + (1 - \lambda^*)\left(\int_{\lambda^*}^1\lambda\frac{1}{1 - \lambda^*}\,\mathrm{d}\lambda\right)^2\right) \tag{19}$$

Simplifying (19) we get:

$$W_M = \mu\left(\frac{1}{\lambda^*}\int_0^{\lambda^*}\lambda^2\,\mathrm{d}\lambda + \frac{1}{1 - \lambda^*}\int_{\lambda^*}^1\lambda^2\,\mathrm{d}\lambda\right) = \frac{\mu}{4}(1 + \lambda^* - \lambda^{*2}) \tag{20}$$

We can write the total welfare as:

$$U = 2\lambda^* (1-\lambda^*)(-c+\theta-\theta+\kappa) + 2(1-\lambda^*)^2 \gamma + \lambda^* \int_0^{\lambda^*} -\frac{\lambda}{\lambda^*} \mathrm{d}\lambda + W_M$$
$$= \frac{1}{4}\Big((8c+8\gamma-8\kappa-\mu-2)\lambda^{*2} + (-8c-16\gamma+8\kappa+\mu)\lambda^* + (8\gamma+\mu)\Big)$$

This is a quadratic equation with respect to λ^*, so when the term with λ^{*2} , $8c+8\gamma-8\kappa-\mu-2<0$, the expression will be maximized at the stationary point

$$\lambda_u^* = \frac{8c+16\gamma-8\kappa-\mu}{2(8c+8\gamma-8\kappa-\mu-2)} \tag{21}$$

$\lambda_u^* \in [0,1]$ when:

$$8c+16\gamma-8\kappa-\mu<0 \wedge 2(8c+8\gamma-8\kappa-\mu-2) < 8c+16\gamma-8\kappa-\mu \Leftrightarrow$$
$$8(\gamma+c-\kappa) < \mu-8\gamma \wedge 8(\gamma+c-\kappa) < \mu+4+8\gamma \Rightarrow$$
$$2\gamma+c-\kappa < \frac{\mu}{8}$$

Thus, the stationary point λ_u^* will be the interior maximizer when: $2\gamma+c-\kappa<\frac{\mu}{8}$. When $2\gamma+c-\kappa>\frac{\mu}{8}\wedge 8c+8\gamma-8\kappa-\mu-2<0$, expression (21) will be negative, and the feasible λ^* that maximizes welfare will the left most feasible λ_u^* (the case that maximizes the number of cooperators). This implies $\theta_u^*=1$.

When $8c+8\gamma-8\kappa-\mu-2>0$, welfare will be maximized at the corner; again when λ_u^* takes the lowest possible value - implying $\theta_u^*=1$. So, for $2\gamma+c-\kappa<\frac{\mu}{8}$, $\lambda_u^*\in[0,1]$, and is defined with expression (21). The corresponding θ_u^* is defined then with (4. Solving for θ_u^* we get: $\theta_u^* = \frac{3(4\gamma+1)(c+\gamma-\kappa)}{-8c-8\gamma+8\kappa+\mu+2} + \frac{1}{2}(c+3\gamma+\kappa) - \frac{\mu}{4}$. For $2\gamma+c-\kappa \geq \frac{\mu}{8}$ the welfare is maximized when $\theta=1$.

□

Proof of Proposition 7: To determine preferred θ for every type i, we shall first calculate θ_i^C and θ_i^D. The payoff of cooperator as a function of θ is given equation (7). Payoff in (7) is strictly concave function of θ. Maximizing with respect to θ we get that when

$$\theta_i^C = \frac{1}{4}(2c-\lambda_i\mu+2\kappa) \tag{22}$$

and payoff of type i when $\theta=\theta_i^C$ is:

$$\pi_i(\theta_i^C) = -\frac{-4\lambda\mu(3c+4\gamma-3\kappa)+4(c-\kappa)^2+8\gamma(\mu+2)+\lambda(\lambda+4)\mu^2+8\lambda\mu}{16(c+\gamma-\kappa-1)-8\mu}$$

Which for the median type has the value:

$$\pi_i(\theta_M^C) = -\frac{(-4c+4\kappa+3\mu)^2+16(4\gamma+\mu)}{64(c+\gamma-\kappa-1)-32\mu}$$

From (8) we have that the preferred level of enforcement institution quality - given that an agent will defect is:

$$\theta_i^D = \frac{1}{2}(c+\kappa-1) - \frac{1}{4}(\lambda_i+1)\mu \tag{23}$$

and the payoff at this level of enforcement institution for a defector is:

$$\pi_i(\theta_i^D) = \frac{-4\mu(2\gamma\lambda + c(\lambda+1) - \kappa(\lambda+1) + \lambda - 1)}{16(c+\gamma-\kappa-1) - 8\mu} + \frac{16\lambda(c+\gamma-\kappa-1) + 4(-c+\kappa+1)^2 + (\lambda+1)^2\mu^2}{16(c+\gamma-\kappa-1) - 8\mu}$$

Which for the median voter takes a value:

$$\frac{8\mu(3c + 2\gamma - 3\kappa - 1) - 16\left((c-\kappa)^2 + 2\gamma - 1\right) - 9\mu^2}{64(c+\gamma-\kappa-1) - 32\mu}$$

The difference of the median voter's payoffs at the preferred levels of θ conditional on action in the first stage is:

$$\pi_i(\theta_M^C) - \pi(\theta_M^D) = \frac{(2\gamma+1)(\mu+2)}{-8c - 8\gamma + 8\kappa + 4\mu + 8} \tag{24}$$

The numerator in (24) is positive, and as for the denominator, $-8c - 8\gamma + 8\kappa + 4\mu + 8 > 0 \Leftrightarrow \gamma + c - \kappa < 1 + \frac{\mu}{2}$. And this is the condition stated in (3). As we are interested only in region of parameters where there is unique equilibrium, $\pi_i(\theta_M^C) - \pi(\theta_M^D)$ will always be positive. The preferred level of enforcement θ for the median voter will thus be: $\frac{c+\kappa}{2} - \frac{\mu}{8}$. $\square$

References

Benabou, R. and Tirole, J. (2006). Incentives and prosocial behavior. *American Economic Review*, 96(5):1652–1678.

Benabou, R. and Tirole, J. (2011). Laws and norms. Technical report, National Bureau of Economic Research.

Bergstrom, T. C. (2003). The algebra of assortative encounters and the evolution of cooperation. *International Game Theory Review*, 5(03):211–228.

Bulow, J. I., Geanakoplos, J. D., and Klemperer, P. D. (1985). Multimarket oligopoly: Strategic substitutes and complements. *The Journal of Political Economy*, pages 488–511.

Dhillon, A. and Rigolini, J. (2011). Development and the interaction of enforcement institutions. *Journal of Public Economics*, 95(1):79–87.

Dixit, A. (2003). Trade expansion and contract enforcement. *Journal of Political Economy*, 111(6):1293–1317.

Dixit, A. (2009). Governance institutions and economic activity. *American Economic Review*, 99(1):5–24.

Fehr, E., Klein, A., and Schmidt, K. M. (2001). Fairness, incentives and contractual incompleteness.

Gneezy, U. and Rustichini, A. (2000). Fine is a price, a. *J. Legal Stud.*, 29:1.

Hoppe, H. C., Moldovanu, B., and Sela, A. (2009). The theory of assortative matching based on costly signals. *The Review of Economic Studies*, 76(1):253–281.

Jewitt, I. (2004). Notes on the shape of distributions. Technical report, Mimeo, Oxford University, Summer.

Kranton, R. E. (1996). Reciprocal exchange: a self-sustaining system. *The American Economic Review*, pages 830–851.

Levy, G. and Razin, R. (2013). Calvin's reformation in geneva: self and social signalling. *Journal of Public Economic Theory*.

McAfee, R. P. (2002). Coarse Matching. *Econometrica*, 70(5):2025–2034.

Ok, E. A. and Vega-Redondo, F. (2001). On the evolution of individualistic preferences: An incomplete information scenario. *Journal of Economic Theory*, 97(2):231–254.

Seabright, P. (2009). Continuous Preferences and Discontinuous Choices: How Altruists Respond to Incentives. *The B.E. Journal of Theoretical Economics*, 9(1):1–28.

Titmuss, R. M. (1970). The gift relationship: from human blood to social policy.

Vives, X. (2000). *Oligopoly pricing*. MIT Press, Cambridge, Mass. [u.a.].