

# The Scope of Cooperation: values and incentives\*

Guido Tabellini

IGIER, Bocconi University; CEPR; CES-Ifo; CIFAR

First version: January 2007; this version: February 2008.

## Abstract

What explains the range of situations in which individuals cooperate? This paper studies a model where individuals respond to incentives but are also influenced by norms of good conduct inherited from earlier generations. Parents rationally choose what values to transmit to their offspring, and this choice is influenced by the spatial patterns of external enforcement and of likely future transactions. The equilibrium displays strategic complementarities between values and current behavior, which reinforce the effects of changes in the external environment. Values evolve gradually over time, and if the quality of legal enforcement is chosen under majority rule, there is path dependence: adverse initial conditions may lead to a unique equilibrium where legal enforcement remains weak and individual values discourage cooperation.

---

\*Forthcoming *Quarterly Journal of Economics*. I thank Roberto Barro, Alberto Bisin, Ed Glaeser, Avner Grief, Richard Holden, three anonymous referees and seminar participants at Brown University, CIFAR, IGIER, the IMF research department, Kellogg, the NBER Summer Institute, NYU and the University of Naples for their comments, and Carlo Prato for excellent research assistance. I am grateful to CIFAR and Bocconi University for financial support.

# 1 Introduction

What determines the range of situations in which individuals cooperate? This question has been addressed by a large literature in economics, political science and sociology. The traditional approach by economists poses this question in terms of reputation: the scope of cooperation is explained by the strength of the incentives to preserve one's reputation in repeated interactions, relative to the temptation to cheat.<sup>1</sup>

While the traditional economic approach has yielded important insights, it misses an important dimension. In many social situations individuals behave contrary to their immediate material self interest, not because of an intertemporal calculus of benefits and costs, but because they have internalized a norm of good conduct. Whether we risk our lives fighting in war, or bear the cost of voting in large elections, or refrain from stealing or cheating in an economic transaction, is also determined by our values and beliefs about what is right or wrong.<sup>2</sup> This observation raises several natural questions: what is the origin of specific norms of good conduct? What determines the range of situations over which they are meant to apply? Why do specific values persist in some environments and not in others? How do values evolve over time? And how do they interact with economic incentives, and with the economic and political environment?

Until recently and with few exceptions, economists have refrained from asking these questions and have accepted a division of labor. Other social sciences, primarily sociology, discuss the endogenous evolution of values and preferences. Economics studies the effects of incentives on individual decisions and aggregate outcomes, taking individual preferences as given. Even when social norms have been acknowledged as playing a crucial role, as in the selection of focal points when there are multiple equilibria, economists have studied the implications of these norms, but not their endogenous evolution. A byproduct of this division of labor is that, until recently, the analysis of

---

<sup>1</sup>Dixit (2004) provides an excellent overview and makes several original contributions taking the economic approach. Axelrod (1984) and Gambetta (1988) are influential contributions in political science and sociology, that overlap with the economic approach.

<sup>2</sup> A large literature in the natural sciences and evolutionary psychology discusses the role of emotions in regulating and motivating human behavior, suggesting an evolutionary explanation of our moral capacities. See for instance Barkow, Cosmides and Tooby(1992), Pinker (1997), Massey (2002) and other references quoted in Kaplow and Shavell (2007). See also the evidence in Fehr, Fischbacher and Gächter (2002).

social norms has generally escaped the discipline of methodological individualism, the paradigm of economics. This is unfortunate because, as stressed by Kaplow and Shavell (2007), values and moral rules are malleable and entail an element of rational choice. The principles that children learn in the family or at school, or the codes of conduct that regulate adult individual behavior, are the outcome of purposeful choices and rational deliberations. As such, they lend themselves to be studied with the traditional methods of economics.<sup>3</sup>

This paper studies the scope of cooperation combining ideas from economics and sociology. Throughout I neglect the role of reputation, and view cooperation as resulting from a tradeoff between material incentives and individual values. From sociology I borrow the question and the emphasis on norms of good conduct. Namely, I ask how individual values that sustain cooperation evolve endogenously over time. But I address this question with the traditional tool kit of economists, individual optimization and equilibrium analysis, and I focus on how values interact with economic incentives.

The model is adapted from Dixit (2004). Individuals are randomly matched with others located along a circle, to play a prisoner's dilemma game. But unlike in Dixit (2004), they play only once, so there is no role for reputation, and cooperation is sustained by individual values (a pleasure from cooperating with nearby individuals). The scope of cooperation corresponds to the set of matches over which cooperation is sustained, and this depends both on economic incentives and individual values.

The model is designed to capture an important idea stressed by sociol-

---

<sup>3</sup>Besides the pathbreaking work of Gary Becker (see Becker 1993, 1996), recent contributions by economists have started undermining this division of labor. Guiso et al. (2006) and Fernandez (2007a, 2008) discuss much of this recent (mainly empirical) literature, while Ben-Ner and Putterman (1998) discuss earlier contributions. See also Algan and Cahuc (2006), (2007), Barro and McCleary (2006) and Giuliano (2007). Important theoretical contributions include Bisin and Verdier (2001), Akerlof and Kranton (2000), (2006), Benabou and Tirole (2006a). Bisin and Verdier (2008) provide a review. Hauk and Saez-Marti (2002), Francois and Zabojnik (2005) and Francois (2006) have applied the methodology pioneered by Bisin and Verdier (2001) to closely related issues. Other recent contributions by economists, with a similar approach but a different focus, include Benabou and Tirole (2006b), Lindbeck and Nyberg (2006), Doepke and Zilibotti (2005). The literature by sociologists on these issues is just too large to be mentioned here. But see Nisbet and Cohen (1996) for an interesting example of an attempt to explain the endogenous evolution of individual values. Bowles and Gintis (2002) have also formally studied the evolution of norms facilitating cooperation, but without relying on individual optimization.

ogists, that rests on the distinction between limited vs generalized morality (Banfield 1958, Platteau 2000). Norms of limited morality are applicable only to a narrow circle of friends or relatives; with others, cheating is allowed and regularly occurs. Generalized morality instead applies generally towards everyone, and entails respect for abstract individuals and their rights. Individuals who have internalized norms of generalized morality cooperate over a larger range of situations.

To analyze how such values evolve endogenously, I build on the work of Bisin and Verdier (2001), Bisin, Topa and Verdier (2004). Parents optimally choose what values to pass on to their children, but evaluate their children's welfare with their own values. This assumption of "imperfect empathy" implies that the equilibrium is both forward and backward looking. It is backward looking, because the parents' values influence their educational choices. Thus, values evolve gradually over time and during the transition they reflect historical features of the external environment. But the equilibrium is also forward looking, since parents adapt their educational choices to the future environment of their children. This creates a strategic complementarity between values and behavior. If more individuals follow a norm of generalized morality, then those who abide by this norm expand the scope of cooperation. And conversely, an expansion in the scope of cooperation facilitates the diffusion of norms of generalized morality. Thus, values and behavior mutually reinforce each other, and this strengthens the effects of changes in the environment.

In equilibrium, the diffusion of values reflects the spatial pattern of external enforcement and of likely future transactions. Generalized morality is hurt if external enforcement of cooperation is mainly local, while it is encouraged by strong enforcement of more distant transactions. Thus, well functioning legal institutions breed good values, since legal enforcement is particularly relevant between unrelated individuals. Instead, better informal enforcement sustained by ongoing relations in a closed network may be counterproductive for values. This conclusion is reinforced by the effects of the matching technology: localization of economic activity hurts the diffusion of values that sustain generalized cooperation. Thus, the model predicts that clan based societies develop very different value systems compared to modern societies that rely on the abstract rule of law. But extreme globalization can also be detrimental to values, if it increases the likelihood of situations where the parents' codes of good conduct are not applicable or have weak implications.

Values do not only matter in the individual decision to cooperate in bilateral transactions. They also influence collective decisions. If legal enforcement is set in a political equilibrium, then there is path dependence. When limited morality initially prevails, the economy ends up in a steady state with lax legal enforcement, poor values and lack of cooperation. And viceversa, if a norm of generalized morality is initially widespread, then the equilibrium converges to a steady state with opposite features.<sup>4</sup>

These results can explain the puzzling persistence of institutional outcomes emphasized by Acemoglu et al. (2001), Acemoglu and Robinson (2006), Glaeser et al. (2004) amongst others. Countries and regions that centuries ago were ruled by despotic governments, or where powerful élites exploited uneducated peasants or slaves, today are plagued by institutional and organizational failures. They also display limited morality and lack of trust between individuals (Banfield 1958, Putnam 1993, Guiso et al. 2008, Tabellini 2005, 2008). The idea that culture is the missing link between distant history and current institutional performance is also supported by the attitudes of 2nd generation US citizens: their trust is higher if they came from countries that over a century ago had better political institutions - Tabellini (2008).

This lack of social capital in environments with a history of political abuse and exploitation could be both an independent cause and an effect of current institutional failures. But this paper suggests that it is bound to be very difficult to pin down which specific institutional features are responsible for observed outcomes. In equilibrium, the form and functioning of institutions are jointly determined with the value systems, and their evolution is dictated by initial and possibly random historical circumstances.

The outline of the paper is as follows. Section 2 outlines the model in its simplest version with exogenous preferences for cooperation. Section 3 makes preferences endogenous and shaped by the educational choices of optimizing parents. Section 4 adds politics and studies the equilibrium with endogenous preferences and endogenous policies. Section 5 discusses some extensions.

---

<sup>4</sup>The two way interactions between the external environment and endogenous beliefs or values has also been studied by Benabou and Tirole (2006a), Bisin and Verdier (2000, 2004), Francois (2006) and Hauk and Saez-Marti (2002) in different settings.

## 2 The Scope of Cooperation with Exogenous Values

### 2.1 The Model

The model is adapted from Dixit (2004), chapter 3. A continuum of one-period lived individuals is uniformly distributed on the circumference of a circle of size  $2S$ . Thus the maximum distance between two individuals is  $S$ . The density of individuals per unit of arc length is 1. A larger circle (a higher  $S$ ) implies a lower population density and thus a more heterogeneous society along the relevant spatial dimension.

Each individual is randomly matched with another located at distance  $y$  with probability  $g(y) > 0$ . Only distance matters, not location, and no restriction is placed on  $g(\cdot)$  except that the probabilities of all matches between 0 and  $S$  sum to 1,  $\int_0^S g(y) = 1$ . Distance could refer to geography, but also to social or economic dimensions such as religion, ethnicity, class, and so on.

The two matched individuals observe their distance and play a prisoner's dilemma game. Each player simultaneously chooses whether to cooperate ( $C$ ) or not to cooperate ( $NC$ ). Their material payoffs are:

**Table 1. Payoffs from Prisoner's Dilemma**

	$C$	$NC$
$C$	$c, c$	$h - l, c + w$
$NC$	$c + w, h - l$	$h, h$

As in any prisoner's dilemma game,  $c > h$  and  $l, w > 0$ . I also assume throughout that  $l \geq w$ , namely that the loss of being cheated is at least as large as the benefit of cheating (relative to the respective payoffs under full cooperation and no-cooperation at all); as shown in a previous version and further discussed below, this assumption rules out asymmetric equilibria where identical players choose opposite strategies.

Besides these material payoffs, each individual enjoys a non-economic (psychological) benefit  $d > \max(l, w)$  whenever it plays  $C$ , irrespective of how his opponent played. This non-economic benefit decays with distance at exponential rate  $\theta > 0$ . Thus, playing  $C$  against an opponent located at distance  $y$  results in non-economic benefit  $d e^{-\theta y}$ . Since the payoffs of the prisoner's dilemma game are the same for all matches, the parameter  $\theta$  can be

interpreted as the rate at which non-economic benefits decay with distance, relative to the economic payoffs.<sup>5</sup>

This formulation captures two plausible ideas. First, individuals are motivated by more than just material payoffs. They also value the act of cooperating per se. These “warm glow preferences” (Andreoni 1990) could reflect religious or moral principles, or other values that induce individual self regulation in social interactions. Second, these norms of good conduct apply with particular force within a circle of socially connected individuals, but they are weaker with less familiar individuals. This trait might have evolved from a distant past when social interactions were mainly confined to small groups of hunters-gatherers (Cosmides and Tooby 1992). Or it could be induced by an ability to detect true motives and character through frequent interactions, which would encourage honest non-instrumental behavior with more familiar people (Cooter 1998). Below I discuss a large body of evidence consistent with the idea that moral ties are strengthened by familiarity.

Finally, I assume that there are two types of player indexed by  $k = 0, 1$ . Both types enjoy the same benefit  $d$  of cooperating, but they differ in the rate at which the benefit decays with distance,  $\theta^1$  and  $\theta^0$ , with  $\theta^0 > \theta^1$ . For shortness, I call them trustworthy or “good” (if  $k = 1$ ) and not-trustworthy or “bad” (if  $k = 0$ ), since for any positive distance a good player values cooperation more than a bad player. Individuals observe distance,  $y$ , but not the trustworthiness of their partner. The fraction of good ( $k = 1$ ) types in the population is the same at any point in the circle, and in this section it is a fixed parameter  $n$ , with  $1 > n > 0$ . The next section makes  $n$  endogenous.

## 2.2 Equilibrium

Consider a player deciding whether to play  $NC$  or  $C$  in a match of distance  $y$ . Let  $\pi(y)$  denote the probability that his/her opponent will play  $C$ . We can express the player’s net expected material gain from playing  $NC$  rather than  $C$  as:

$$T(\pi(y)) = [l - \pi(y)(l - w)] > 0 \tag{1}$$

---

<sup>5</sup>A previous version (CEPR Discussion Paper 6534) discussed an extension that allowed for reciprocity in the non-economic benefit of cooperation. Namely, the benefit  $d e^{-\theta y}$  is obtained only if the opponent cooperates, but not if he cheats. All the results discussed in the paper go through, except that reciprocity induces additional strategic complementarities and hence additional equilibria.

This expression captures the temptation not to cooperate. It is strictly positive, as it is always better not to cooperate.

Given the assumption that  $l \geq w$ , the function  $T(\pi(y))$  is non-increasing in  $\pi(y)$ , the probability that the opponent will play  $C$ , and strictly decreasing in  $\pi(y)$  if  $l > w$ . Intuitively, if  $l > w$ , then the loss of being cheated is greater than the benefit of cheating (relative to the respective payoffs under full cooperation and no-cooperation at all). This means that the temptation to cheat is greater if the opponent is also more likely to cheat (i.e.  $T(\cdot)$  is strictly decreasing in  $\pi$ ). Thus, the game entails a strategic complementarity.<sup>6</sup>

This temptation is the same for all players. It must be balanced against the non-economic benefit of cooperation,  $de^{-\theta^k y}$ , which instead depends on a player's type. A type  $k = 0, 1$  is just indifferent between playing  $C$  or  $NC$  in a match with someone at distance  $\tilde{y}^k$  if:

$$T(\pi(\tilde{y}^k)) = de^{-\theta^k \tilde{y}^k} \quad (2)$$

Solving for  $\tilde{y}^k$ , we obtain:

$$\tilde{y}^k = \{ \ln d - \ln [(w - l) \pi(\tilde{y}^k) + l] \} / \theta^k \quad (3)$$

Note that the benefit of cooperation,  $de^{-\theta^k y}$ , is strictly decreasing in  $y$ . This follows from the assumption that the norm of good conduct applies with greater strength to closer partners. Hence, holding  $\pi$  constant, this individual prefers to play  $C$  in a match with someone at distance  $y < \tilde{y}^k$ , and she/he prefers to play  $NC$  if  $y > \tilde{y}^k$ .

To characterize the equilibrium, we need to pin down the equilibrium probability of cooperation  $\pi(y)$  for all possible values of  $y$ . If  $l > w$ , then the right hand side of (3) is increasing in  $\pi(y)$ , and there are multiple equilibria. Here we confine attention to the Pareto superior equilibrium that sustains maximal cooperation (a previous version, CEPR Discussion Paper 6534, characterized all equilibria). In such equilibrium, players entertain the most optimistic expectations about their opponent's behavior, consistent with incentive compatibility.

Consider a bad player,  $k = 0$ , and suppose that she/he expects the opponent always to cooperate, so that  $\pi(y) = 1$ . Then (3) reduces to:

$$Y^0 = [\ln d - \ln w] / \theta^0 \quad (4)$$

---

<sup>6</sup>If, contrary to our assumption,  $w > l$ , then the actions would be strategic substitutes.



By construction, player  $k = 0$  finds it optimal to cooperate up to distance  $y \leq Y^0$  and to play  $NC$  if  $y > Y^0$ .

What about a good player,  $k = 1$ ? Up to distance  $y \leq Y^0$  a good player also plays  $C$ , since he/she faces the same temptation as a bad player but has a higher non-economic benefit from cooperation. For  $y > Y^0$ , the good player realizes that all the bad players will play  $NC$ , and the good player's most optimistic expectation is that their opponent cooperates only if he/she is good, which happens with probability  $n$ . Inserting  $\pi(y) = n$  on the right hand side of (3) we thus obtain:

$$\tilde{y}^1 = [\ln d - \ln [(w - l)n + l]] / \theta^1 \quad (5)$$

If  $\tilde{y}^1 > Y^0$ , then by construction a good player finds it optimal to cooperate up to  $\tilde{y}^1$ , given that he/she expects all other good players to also cooperate; beyond this threshold the good player prefers to play  $NC$ . If instead  $\tilde{y}^1 \leq Y^0$ , then (given the expectation that everyone cooperates up to  $Y^0$ ), the good player also prefers to play  $C$  up to  $Y^0$ , but not beyond. Thus, in equilibrium the upper threshold of cooperation for a good player is:

$$Y^1 = \text{Max} \{ \tilde{y}^1, Y^0 \} \quad (6)$$

Note that  $Y^1 \geq Y^0$ , with strict inequality if  $n$  is sufficiently large, or if the two types are sufficiently different from each other. In particular,  $Y^1 > Y^0$  even at  $n = 0$  if the following condition is satisfied:

$$\frac{\theta^0}{\theta^1} > \frac{\ln(l/d)}{\ln(w/d)} \quad (A0)$$

Since  $l \geq w$ , this condition implies  $\theta^0 > \theta^1$ , but it is stronger. To reduce the number of possible cases we need to keep track of, from here on I assume that (A0) always holds.<sup>7</sup>

Finally, if  $l = w$ , then right hand side of (3) does not depend on  $\pi(\cdot)$  and the equilibrium is unique.

I summarise this discussion in the following:

**Proposition 1** (i) *In the Pareto superior equilibrium, a player of type  $k$  cooperates in a match of distance  $y \leq Y^k$  and does not cooperate if  $y > Y^k$ , where  $Y^k$  is defined by (4)-(6), for  $k = 0, 1$ . (ii)  $Y^1 \geq Y^0$ , with strict inequality if (A0) holds. (iii) If  $l = w$ , the equilibrium is unique.*

---

<sup>7</sup>This entails no loss of generality. All the results stated below continue to hold if this condition is violated, but the proofs are more complicated as we need to go through more possible equilibrium configurations.

## 2.3 Discussion

This equilibrium provides a simple theory of the scope of cooperation. The first and most straightforward implication is that cooperation is easier to sustain if individuals are close to each other than if they are far apart - whatever the dimension over which distance is measured. This is the same result obtained by Dixit (2003, 2004), but the reason is different. In Dixit, incentives to maintain a reputation are stronger for nearby matches, because information about cheating is more likely to reach relevant future partners. Here instead the result follows almost directly from the assumption that norms of good conduct apply with greater force between closer individuals.

This behavioral implication is consistent with the evidence in Alesina and La Ferrara (2000, 2002), who find that individuals distrust those that are dissimilar from themselves, and that participation in social activities is lower in US localities that are more fragmented (ethnically or economically). This second finding is a direct implication of this model: a larger  $S$  (a more heterogeneous community) implies that on average matches entail a larger distance, and hence less frequent cooperation.

A large body of experimental evidence also supports the prediction that cooperation is easier to sustain between socially closer individuals. In a trust game played by Harvard undergraduates, Glaeser et al. (2000) find that social connectedness increases both the amount sent and returned, suggesting greater trust and trustworthiness. This could reflect the expectation of future repeated interactions outside of the experiment (i.e. the reputational incentives studied by Dixit). But the amount returned is greater between players of the same race and nationality, even after controlling for social connectedness. Moreover, similar results have been obtained in other experiments, where players could not fully identify themselves. In particular, Bohnet and Frey (1999) study a dictator game, and find that one-way identifiability increases the fairness of the outcome. Similarly, Ichino et al. (2007) ran a repeated version of the trust game with PhD students of EUI in Florence, where players know the nationality but not the identity of their partner. They find that Northern Europeans display more trust and trustworthiness between themselves than towards nationals of other European regions.<sup>8</sup> Finally, Dawes and Thaler (1988) summarise several experiments

---

<sup>8</sup>Contrary to these results, however, Fershtman and Gneezy (2001) do not find evidence of a bias towards members of their own group in anonymous versions of the trust and dictator games played by Eastern and Ashkenazic jews in Israel.

where inducing identification with the group (through pre-play discussion or in other ways) increases the amount contributed to public good provision.

In the introductory section we stressed the distinction between limited vs generalized morality, namely between norms of good conduct that apply in a narrow or in a large set of social interactions. The equilibrium above provides an analytical foundation to this distinction. The distance  $Y^0$  can be interpreted as a small familiar group within which everyone behaves well. Interactions beyond  $Y^0$  refer to the market or to a larger and more anonymous set of individuals, where not everyone can be trusted. The scope of maximal sustainable cooperation over these more distant matches is defined by  $Y^1$ . In principle global cooperation anywhere in the circle could be sustained, but we implicitly assume that  $Y^1 < S$ , where  $2S$  is the size of the circumference.<sup>9</sup>

The variables  $Y^0$  and  $Y^1$  summarize all the relevant determinants of the scope of cooperation, in nearby or more distant matches. In particular, the lower threshold  $Y^0$  rises if the benefit of cheating ( $w$ ) falls, if the non-economic benefit of cooperation ( $d$ ) rises, and if norms of good conduct decay more slowly with distance (if  $\theta$  falls). The upper threshold  $Y^1$  depends on these same variables, but it is also decreasing in the loss from being cheated ( $l$ ) and, if  $l > w$ , it is increasing in the fraction of good players,  $n$ . This follows from imperfect information: as individuals cannot observe their opponent type, in equilibrium the good players bear the risk of cooperating against a cheating opponent. The smaller is the resulting loss, the larger is the range of matches over which cooperation can be sustained. The effect of  $n$  on  $Y^1$  reflects the strategic complementarity in the prisoner's dilemma game: given  $l > w$ , individuals are more willing to cooperate the higher is the probability that their partner will also cooperate.<sup>10</sup>

As discussed more at length in the next sections, the parameters  $w$  and  $l$  in the prisoner's dilemma game reflect the quality of institutions responsible for external enforcement: better enforcement corresponds to a smaller benefit of cheating and a smaller loss from being cheated, and hence an enlarged scope of cooperation (larger  $Y^0$  and  $Y^1$ ). In this model with exogenous pref-

---

<sup>9</sup>Evaluating (5) at  $n = 1$ , this corresponds to assuming:  $S > \frac{1}{\theta r} \ln(d/w)$

<sup>10</sup>These results are similar to those obtained by Dixit (2004) in his model based on reputation, despite the different reason why here individuals refrain from cheating. In contrast to Dixit (2004), however, here the range of cooperation does not depend on the matching technology  $g(y)$ , nor on the overall size of the economy,  $S$ . Baron (2007) considers various extensions of the model of this section (with exogenously given parameters  $\theta^1$  and  $\theta^0$ ) and obtains additional interesting results.

erences there is no dynamics, however. As the external environment changes, individuals react immediately by altering their equilibrium behavior. What matters is current enforcement, not institutions in the distant past. Hence, this version of the model is unable to explain institutional persistence.

Finally, note that neither  $Y^0$  nor  $Y^1$  depend on the parameters  $c$  and  $h$  in the Prisoner's dilemma game (the payoffs in the symmetric outcomes of full cooperation and no-cooperation respectively). The reason is that equilibrium strategies are derived by comparing the payoffs under cooperation vs no-cooperation, given the expected behavior of the opponent - see (1). Hence only relative payoffs influence individual behavior.

### 3 Endogenous Values

#### 3.1 The model

This section studies the endogenous evolution of the distribution of types,  $k = 0, 1$ . Our goal is to study how parents rationally choose what values to transmit to their children, and how this choice is affected by economic incentives and by features of the external environment. Whether a given individual is of type  $k = 0$  or 1 reflects two forces: the exogenous influence of nature or of the external environment, and the deliberate and rational efforts of parents, through education or time spent with their children. Parents are altruistic and care about the utility of their offspring, but evaluate their kid's expected welfare with their own preferences. This assumption of "imperfect empathy" (cf. Bisin and Verdier 2001) implies that some parents devote effort to try and shape the values of their children to resemble their own.

Specifically, consider an ongoing economy that lasts forever. Individuals live two periods. In the first period of their life they are educated by their parents and, once education is completed, they are active players in the game described above. In the second period, each individual is the parent of a single kid and the parent's only activity is to devote effort to educate his/her kid. Parental education increases the probability that the kid becomes good (i.e. that  $\theta^k = \theta^1$ ), but it is costly for the parent. To obtain a closed form solution we assume a quadratic cost function:  $\frac{1}{2\varphi}f^2$ , where  $f \geq 0$  denotes parental effort, and  $\varphi > 0$  is a parameter that captures the rate at which the marginal cost increases with effort (lower  $\varphi$  corresponds to a marginal cost that rises more rapidly). Effort is chosen by each parent before observing their kid's

value. Conditional upon effort, the probability of having a good kid does not depend on the parent's type. Specifically, given effort  $f \geq 0$ , the kid turns out to be good ( $\theta^k = \theta^1$ ) with probability  $\delta + f$ , and bad ( $\theta^k = \theta^0$ ) with probability  $1 - \delta - f$ , where  $1 > \delta > 0$  reflects the choice by nature.<sup>11</sup>

Once parents have completed the education, each young player observes his/her own type and plays the matching game described in the previous section. Thus, in any period  $t$  the economy behaves exactly like in the previous section, except that here the composition of types is endogenous and varies with time. Throughout, I restrict attention to the Pareto superior equilibrium of the matching game, as described in Proposition 1.

Consider a parent of type  $p$  who has a kid of type  $k$  in period  $t$ , for  $k, p = 0, 1$ . Let  $V_t^{pk}$  denote the parent's evaluation of their kid's overall expected utility in the equilibrium of the matching game. Recalling that the probability of a match with someone located at distance  $z$  is denoted  $g(z)$ , we have:

$$V_t^{pk} = U_t^k + d \int_0^{Y_t^k} e^{-\theta^p z} g(z) dz \quad (7)$$

The first term  $U_t^k = U(\theta^k, n_t)$  denotes the expected equilibrium material payoffs of a kid of type  $k$ , in a game with a fraction  $n_t$  of good players. An expression for  $U_t^k$  is provided in subsection 1 of the appendix. The second term on the RHS of (7) is the parent's evaluation of their kid's expected non-economic benefit of cooperating in matches of distance smaller than  $Y_t^k$ . Note that this evaluation is done with the parent's value parameter,  $\theta^p$ , rather than with the kid's value. Thus, if the kid is born with the same value of his/her parent (if  $\theta^p = \theta^k$ ), then parent and kid evaluate the outcome of the kid's

---

<sup>11</sup>Note the asymmetry. Parents can exert effort to increase the expected trustworthiness of their kid, but not to reduce it. With a slight change in notation, this asymmetry can be interpreted almost literally as saying that inculcating trustworthiness in one's kid is costly, while inculcating dishonesty or non-trustworthiness does not cost any effort to the parent. A previous version of this paper removed the asymmetry, and assumed that it was equally costly to increase or decrease the trustworthiness of one's kid, relative to the choice made by nature. The qualitative results were similar, although additional conditions on parameter values had to be imposed to obtain some of the comparative statics results discussed below. Unlike in Bisin and Verdier (2001), and given the different focus of the analysis, I neglect the possibility that the kids' values or the effect of parental effort also depend on the current distribution of types in the population. This implies that to obtain dynamic stability we need additional conditions on parameter values.

matching game identically. But if they have different values, then  $V_t^{pk}$  differs from the kid's own evaluation: the value parameter in the last term on the right hand side of (7),  $\theta^p$ , is that of the parent, while the relevant distance threshold according to which the game is played,  $Y_t^k$ , is that of the kid.

Why should parents impose their own values in the evaluation of their kid's welfare, rather than using the kid's preferences? If preferences refer to the evaluation of alternative material payoffs, this view is not very compelling. There is no strong reason why a parent should care about whether their kid prefers wine or beer, or whether he/she enjoys rock or classical music. But in the context of (7), parents express a value judgement on their kids' actions, and values are not randomly chosen to suit one's tastes. They reflect deeply held convictions about religious or moral principles, or beliefs about the long run consequences of alternative patterns of behavior that are likely apply to everyone. In other words, values are not the same thing as preferences. Parents are likely to be convinced that what is "right" for themselves is also "right" for everyone else, and in particular for their kids.

This assumption of "imperfect empathy" is not crucial for the results, however. As discussed below, the model implies that in equilibrium the good players enjoy higher overall expected utility (inclusive of the "warm glow" benefit of cooperation) than the bad players, despite their being cheated in some matches. Hence, even a purely altruistic parent, who evaluates their kid's welfare with the kid's own values, would be prepared to bear a cost to increase the probability that their kid is a good type. In this case, however, both good and bad parents would want to transmit similar values, and to induce persistence in the distribution of types one would have to assume that a parent's type somehow influences the effectiveness of their educational effort. This point is further discussed below.<sup>12</sup>

### 3.2 The parent's optimization problem

This subsection describes how parents choose effort,  $f_t$ . I start by showing that a parent always prefers to have a kid with their own values:

**Lemma 2** *If  $k \neq p$ , then  $V_t^{pp} \geq V_t^{pk}$ , with strict inequality if  $Y^1 > Y^0$ .*

---

<sup>12</sup>An editor suggested an additional reason why parents may want to inculcate good values in their children: it makes the task of being a parent much easier !

The proof is in the appendix. This intuitive result reflects two assumptions. First, individual types are not observable, and hence there is no incentive for strategic delegation (i.e. there is no strategic gain in distorting the kid's preferences when the child plays the subsequent game). If types were observable, then a good player would induce his/her opponent to cooperate over a longer distance range if the opponent were also good. Even a bad parent would internalize this benefit, and this might induce a general preference for having a good kid. Second, imperfect empathy implies that the only reason for changing one's kid value  $\theta^k$  is to induce the kid to change her/his behavior. The psychological benefit of cooperation is evaluated by the parent with their own discount rate  $\theta^p$ . Hence the parent does not directly benefit from a lower  $\theta^k$ , except through the induced effects on the kid's behavior.

Note that our maintained assumption (A0) implies  $Y^1 > Y^0$ , and hence any parent strictly prefers to have a kid with their own values.

Given that effort to educate one's kid costs the parent some disutility according to the quadratic function summarized above, and given that parental effort is chosen before observing the kid's type, Lemma 2 immediately implies:

**Corollary 3** *A "good" parent ( $p = 1$ ) exerts strictly positive effort. A "bad" parent ( $p = 0$ ) exerts no effort.*

Intuitively, by Lemma 2, a bad parent would like to have a bad kid. Hence, a bad parent will never exert any effort to increase their kid's expected trustworthiness. Conversely, a good parent would like to have a good kid. Hence at the margin a good parent is prepared to exert at least some effort to increase the probability of this happening.

Given this result, the fraction of good players in period  $t$ ,  $n_t$ , evolves endogenously over time according to:

$$n_t = n_{t-1}(\delta + f_t) + (1 - n_{t-1})\delta = \delta + n_{t-1}f_t \quad (8)$$

where from here onwards, with a slight abuse of notation,  $f_t$  denotes effort by the good parent. Intuitively, if parents exerted no effort, then the average fraction of good kids in the population would just equal  $\delta$ . But the good parents (of which there is a fraction  $n_{t-1}$  in period  $t$ ) exert effort  $f_t$  in period  $t$ , and this increases the fraction of good kids in the population by  $n_{t-1}f_t$  on average.

Next, consider the optimal choice of effort. A good parent takes as given the effort chosen by all other parents, and takes into account the equilibrium implications of his/her own kid's value for his/her own welfare, according to (7). At an interior optimum, the first order optimality condition equates the marginal cost and the expected net marginal benefit of effort, and by (7) it can be written as:

$$f_t/\varphi = (U_t^1 - U_t^0) + d \int_{Y^0}^{Y_t^1} e^{-\theta^1 z} g(z) dz \quad (9)$$

Consider the right hand side of (9), that captures the net marginal benefit of effort. The first term is the change in the kid's expected material payoffs, if his/her value switches from  $\theta^0$  to  $\theta^1$ . This term is always negative, since for any probability that the partner in a match will cooperate, the kid's expected material payoffs are always higher if the kid plays *NC* (see subsection 2 of the appendix). The second term is the expected benefit of extending the scope of the kid's cooperative behavior to a larger range of matches, evaluated with the parent's values,  $\theta^p = \theta^1$ . This term is always positive, since extending the scope of the kid's cooperative behavior increases the direct non-economic benefit enjoyed by the parent. Hence, the parent perceives a tradeoff. Increasing his/her kid's trustworthiness hurts the kid's expected material payoffs, and this cost is internalized by the parent. But a good kid also provides expected direct non-economic benefits to the parent. By Corollary 3, we know that the benefits exceed the costs, and hence  $f_t > 0$ .

Subsection 2 of the Appendix computes the equilibrium expression for  $U_t^1 - U_t^0$  and shows that the parents' optimality conditions, (9), yields:

$$\begin{aligned} f_t &= \varphi d \left[ -e^{-\theta^1 Y_t^1} \int_{Y^0}^{Y_t^1} g(z) dz + \int_{Y^0}^{Y_t^1} e^{-\theta^1 z} g(z) dz \right] \\ &= \varphi d \left\{ -e^{-\theta^1 Y_t^1} + E[e^{-\theta^1 y} \mid Y_t^1 \geq y \geq Y^0] \right\} \Pr(Y_t^1 \geq y \geq Y^0) \quad (10) \end{aligned}$$

Note that  $\delta + f_t$  denotes a probability and that by (10)  $f_t > 0$ . Thus, implicit in (10) is the restriction that the parameter  $\varphi$  is sufficiently small that  $1 - \delta \geq f_t$  (see also the discussion of dynamic stability below). Equation (10) defines  $f_t$  as a known function of  $Y_t^1$ ,  $f_t = F(Y_t^1)$  - all other terms on the right hand side of (10) are fixed parameters, including  $Y^0$ . Subsection 2 of the appendix proves:



**Lemma 4** *The function  $F(Y_t^1)$  is strictly increasing in  $Y_t^1$ .*

Intuitively, if the difference in behavior between good and bad players increases (as captured by the variable  $Y_t^1$ ), then good parents are induced to put more effort to increase the probability of having a good kid. That is, effort increases as the behavioral implications of the kids' values become more relevant.

This property is important, because it gives rise to a second strategic complementarity. If parents expect others to put more effort into education, they anticipate that the fraction of good players will increase. They realize that this will expand the scope of cooperation,  $Y_t^1$ , and as a result they exert more effort. In fact, it is easy to verify that the educational game described in this section is supermodular (cf. Amir 2003).

### 3.3 The equilibrium

Replacing  $f_t$  with  $F(Y_t^1)$  in (8) and simplifying, the equilibrium is thus given by the vector  $(Y_t^{1*}, n_t^*)$  that solves the following two equations:

$$Y_t^1 = [\ln d - \ln [(w - l)n_t + l]] / \theta^1 \equiv Y(n_t) \quad (11)$$

$$n_t = \delta + n_{t-1}F(Y_t^1) \equiv N(Y_t^1, n_{t-1}) \quad (12)$$

The first equation defines the maximum distance  $Y_t^1$  that sustains cooperation by the good players, as a function of the fraction of other good players in the population,  $Y_t^1 = Y(n_t)$ . As discussed in Section 2,  $Y_t^1$  is a non-decreasing function of  $n_t$ , and it is strictly increasing (and convex) in  $n_t$  if  $l > w$ , that is if the prisoner's dilemma game entails strategic complementarity.

The second equation defines the law of motion of the fraction of good players,  $n_t = N(Y_t^1, n_{t-1})$ . As  $Y_t^1$  increases, good parents are induced to put more effort into changing their kid's value (by Lemma 4, the function  $F(Y_t^1)$  is strictly increasing in  $Y_t^1$ ). Hence, the function  $n_t = N(Y_t^1, n_{t-1})$  is also increasing in  $Y_t^1$ .

Together, equations (11) and (12) implicitly define the equilibrium vector  $(Y_t^{1*}, n_t^*)$  as a function of  $n_{t-1}$  :

$$Y_t^{1*} = G^Y(n_{t-1}) \quad (13)$$

$$n_t^* = G^n(n_{t-1}) \quad (14)$$

These two functions are depicted in Figure 1 (for the case of strategic complementarity in the prisoner's dilemma game,  $l > w$ ). Setting  $n_t = n_{t-1} = n_s$ , we obtain the steady state equilibrium:

$$Y_s^{1*} = Y(n_s^*) \quad (15)$$

$$n_s^* = \frac{\delta}{1 - F(Y_s^{1*})} \quad (16)$$

If  $l > w$ , so that both curves in Figure 1 are upward sloping, multiple equilibria are possible. That is, the same fraction of "good" parents  $n_{t-1}$  might imply more than one equilibrium pair  $(Y_t^{1*}, n_t^*)$ . This reflects the complementarity between values and cooperation in both strategic interactions (amongst parents when choosing education, and amongst kids when playing the matching game). Intuitively, if  $l > w$  then the upper threshold  $Y_t^1$  increases if there are more good players around. This in turn increases the marginal benefit of effort, because differences in behavior between good and bad players become more pronounced. Hence, the payoffs of effort increase with the effort exerted by other parents, which might give rise to multiple equilibria. Since the game is supermodular, however, even with multiple equilibria all comparative statics results apply to the extremal equilibria (Amir 2003).

The equilibrium is unique if the curve  $n_t = N(Y_t^1, n_{t-1})$  always intersects the curve  $Y_t^1 = Y(n_t)$  from left to right, as drawn in Figure 1. Subsection 3 of the appendix proves that a sufficient condition for this to happen is:

$$\frac{1}{\varphi} > l - w \quad (A1)$$

which says that the marginal cost of effort,  $1/\varphi$ , rises sufficiently rapidly, relative to the strategic complementarity captured by  $(l - w)$ . In the remainder of the paper we assume that (A1) holds, so that the equilibrium  $(Y_t^{1*}, n_t^*)$  is unique.<sup>13</sup>

Subsection 3 of the appendix proves that, under condition (A1), the function  $G^n(n_{t-1})$  is strictly increasing in  $n_{t-1}$ , and that  $G^Y(n_{t-1})$  is non-decreasing in  $n_{t-1}$  (strictly increasing if  $l > w$ ). The appendix also proves that there is a  $\bar{\varphi} > 0$  such that, if  $\bar{\varphi} > \varphi > 0$ , then  $dG^n(n_{t-1})/dn_{t-1} < 1$ . We summarize the implications of this discussion in the following:

---

<sup>13</sup>The matching game described in section 2 has multiple equilibria even if (A1) holds, as long as  $l > w$ . But here we are restricting attention to the Pareto superior equilibrium of the matching game.

**Proposition 5** *For  $\varphi > 0$  but small enough, the equilibrium asymptotically reaches the steady state  $(Y_s^{1*}, n_s^*)$  defined by (15)-(16). If  $1/\varphi > l - w$ , then the path towards the steady state is monotonic and the equilibrium  $(Y_t^{1*}, n_t^*)$  is unique. If  $l = w$ , then  $Y_t^{1*}$  is constant and only  $n_t^*$  changes with time. If  $l > w$ , then both  $Y_t^{1*}$  and  $n_t^*$  are time varying and during the adjustment to the steady state they move in the same direction.*

Note that, even with endogenous values, the equilibrium does not depend on the payoffs of the prisoner's dilemma game under full cooperation and no-cooperation ( $c, c$  and  $h, h$  respectively). The reason is that what matters for parental effort is just what happens in the distance range  $Y_t^1 \geq y \geq Y^0$  where the two types behave differently, and over this range the symmetric outcomes of full cooperation and no-cooperation are ruled out.

Finally, how would the equilibrium be affected if the assumption of imperfect empathy was replaced by full altruism (i.e. parents care about the kid's welfare as evaluated by the kids themselves)? To answer, we have to compare the overall expected equilibrium utility of a good vs a bad player. It turns out that a good player is always better off in expected value. Intuitively, the extra psychological benefit obtained when cooperating more than makes up for the lower material payoffs of being cheated. This implies that both good and bad parents, if fully altruistic, would be prepared to devote effort to educate their kid to be good.<sup>14</sup> If the rest of the model is kept as is, then the implications discussed in the next subsection would remain true, except that there would be no dynamics: the steady state is reached after one period, since the initial distribution of player types becomes irrelevant. It would be easy to modify other parts of the model to bring about a slow adjustment to the steady state, however. For instance, one could assume that different parental types face a different cost of educational effort (because education includes setting examples of good behavior), or that the distribution of kids' types also depends on the identity of their parents and not just on their effort

---

<sup>14</sup>The expression for optimal parental effort under full altruism turns out to be the same as the RHS of (10), plus the following positive term, which reflects the difference in psychological benefit between a good and a bad kid in the range where both play  $C$ :

$$\varphi d \int_0^{Y^0} [e^{-\theta^1 z} - e^{-\theta^0 z}] g(z) dz > 0$$

Since this expression does not depend on  $Y^1$ , Lemma 4 would still hold.

(because cultural traits can also be passed across generations inadvertently). Hence, the assumption of imperfect empathy, although plausible in itself, is not the only way to generate the results discussed in the next subsection.

### 3.4 Discussion

We now discuss how the equilibrium is affected by changes in the underlying parameters. Throughout we assume that condition (A1) holds so that the equilibrium is dynamically stable. We also assume that the economy is originally in the steady state,  $(n_s^*, Y_s^{1*})$ .

#### 3.4.1 External enforcement

The payoffs of the prisoner's dilemma game, in particular the temptation to play non-cooperatively and the loss from being cheated, reflect the degree to which cooperation is enforced by external institutions. As already discussed in section 2, an improvement of external enforcement expands the scope of local or general cooperation, as measured by the thresholds  $Y^0$  and  $Y^1$ . But what happens to values, as measured by the fraction of good players,  $n_t$ ? In particular, are good values crowded in or out by better enforcement?

Several papers have asked whether formal and informal institutions are substitutes or complements, focusing on repeated interactions. The general conclusion is that better external enforcement hurts informal institutions because it weakens reputational incentives (Kranton 1996, McMillan and Woodruff 2000, Dhillon and Rigolini 2007). But this issue has generally not been addressed in a setting where cooperation is sustained by values, rather than reputation. Moreover, the claim that formal and informal institutions are mainly substitutes is not very convincing, in light of the observed strong positive correlation between the quality of legal institutions and indicators of trust and trustworthiness (see the evidence in Tabellini 2008).<sup>15</sup>

To address this issue, we model external enforcement as follows. If, in a match of distance  $y$ , one player cooperates and the other does not, then with probability  $q(y)$  cheating goes undetected, and the payoffs are as in Table 1; but with probability  $(1 - q(y))$  cheating is discovered, and playing  $NC$  yields only  $c$  (rather than  $c + w$ ), while playing  $C$  yields  $h$  (rather than  $h - l$ ). If both play  $C$  or  $NC$ , their payoffs are  $c, c$  and  $h, h$  respectively, as in Table

---

<sup>15</sup>A recent exception is Francois (2006), who studies the joint evolution of norms and institutions, and concludes that they are complements.

1. Thus, the quality of external enforcement is captured by the probability of detection,  $1 - q(y)$ , where detection erases the gain from cheating ( $w$ ) and the loss from being cheated ( $l$ ). Better enforcement corresponds to a lower probability  $q(y)$ .

We let the probability of detection depend on distance, because the enforcement method is likely to vary across types of transactions. In general, cooperation is enforced by a variety of means. Some, such as the authority of a local organization or reputation, are relevant for transactions within a narrow group of related individuals. Other, more formal, enforcement methods rely on the law, and are more likely to be relevant in distant transactions where informal enforcement becomes ineffective.

To capture the distinction between informal vs legal enforcement, suppose that  $q(y) = q^0$  for matches of distance  $y \leq \bar{y}$ , where  $\bar{y}$  belongs to a right hand neighborhood of  $Y^0$ . Beyond the distance  $\bar{y}$ ,  $q(y) = q^1$ .<sup>16</sup> Subsection 5 of the appendix shows that, for  $q^0$  and  $q^1$  not very different from each other, all the algebra and the results presented above hold identically, except that the expressions for the thresholds  $Y^k$  are replaced by identical expressions where the parameters  $w$  and  $l$  are replaced by  $\hat{w}^k = wq^k$  and  $\hat{l}^k = lq^k$  respectively). Intuitively, the temptation  $w$  and loss  $l$  are replaced by their expected counterparts.

What are the effects of changes in the enforcement of cooperation in local and distant matches, as captured by the parameters  $q^1$  and  $q^0$ ?

**Better enforcement in distant matches** Consider first the effects of a reduction in the parameter  $q^1$ , taking place at the beginning of period 0, before parents choose effort. This corresponds to an increase of the probability of detection in distant matches, and hence can be interpreted as better enforcement through legal institutions. The algebraic details are in subsection 5 of the appendix, here I discuss the effects referring to Figure 1.

As  $q^1$  is reduced, the curve  $Y_t^1 = Y(n_t)$  in Figure 1 shifts to the right. Intuitively, for a given  $n_t$ , the good players now cooperate over a larger range of matches. Moreover, the threshold  $Y^0$  is not affected by this change. As a result, the curve  $n_t = N(Y_t^1, n_{t-1})$  remains unaffected in period 0. Thus, the scope of cooperation immediately expands ( $Y_t^1$  rises at  $t = 0$ ).

This improvement in the external environment in turn induces parents to

---

<sup>16</sup>Dixit (2004) shows that, if information flows decrease with distance, reputational incentives are stronger for nearby matches than over longer distances.

increase their effort, and  $f_0$  rises - the curve  $N(Y_t^1, n_{t-1})$  is increasing in  $Y_t^1$ , as drawn in Figure 1. Hence, this initial change results in a larger fraction of good players ( $n_0$  rises), which further increases the scope cooperation sustainable in period 0 ( $Y_t^1$  rises further at  $t = 0$ ).

This is not the end of the process, because in period 1 the curve  $N(Y_t^1, n_{t-1})$  shifts upwards. Since more parents are good ( $n_0$  has risen), more of them put effort into educating their children. Hence in period 1 the proportion of good kids is even higher than in period 0 ( $n_1 > n_0$ ) and this brings about an even larger range of cooperative matches in period 1,  $Y_1^1 > Y_0^1$ . The adjustment continues smoothly over time, until a new steady state is reached, with both a larger fraction of good players and a larger scope of cooperation. Thus, better legal enforcement contributes to the diffusion of good values. Because of the strategic complementarities discussed above, cultural forces and economic incentives have self-reinforcing effects. Moreover, a permanent improvement of legal enforcement continues to have effects for many generations after it has occurred, through the educational choices of rational parents. This result is consistent with evidence by Hermann et al. (2008), who study a public good experiment with university students in sixteen countries; they find less pro-social behavior in countries where the rule of law is weaker.

**Better local enforcement** Now consider an improvement of enforcement in nearby matches, that is a lower  $q^0$ . Since this only applies within a narrower range of transactions, it can be interpreted as better enforcement through informal methods, such as reputation or by a local organization. Again, suppose the change takes place at the beginning of period 0, before parents choose effort. Here there is no direct effect on the upper threshold of cooperation  $Y_t^1$ , and hence the curve  $Y_t^1 = Y(n_t)$  does not move in period 0. The lower threshold  $Y^0$  rises, however, as the incentives to cheat in nearby transactions are dampened by better local enforcement. As a result, the difference in behavior between bad and good players shrinks, and parents put less effort into education ( $f_t$  falls in  $t = 0$ , as shown in subsection 5 of the appendix). In terms of Figure 1, the curve  $N(Y_t^1, n_{t-1})$  shifts downwards. This leads to a reduction in the number of good players in period 0 ( $n_0$  falls), and an induced reduction in the upper threshold of cooperation  $Y^1$  in that same period, as the good players now have a weaker incentive to cooperate because there are fewer of them around. From here on, a gradual adjustment to the steady state is set in motion as in the previous case, except that here

both  $Y_t^1$  and  $n_t$  continue to move down (rather than up) for several periods. Thus, better local enforcement destroys good values, and its negative effects are felt over several periods.

**Uniform enforcement** Finally, consider the case in which enforcement is the same for all matches:  $q(y) = q$  for all  $y$ . What are the effects of a uniform improvement of enforcement (a lower  $q$ ), again taking place at the beginning of period 0? Here both curves in Figure 1 shift to the right as enforcement improves. As shown in subsection 5 of the appendix, the net effect on education and hence on values is ambiguous. As a result, over time the fraction of good players  $n_t$  can either rise or fall, setting in motion a transition to a new steady state with higher or lower values of  $(n_s^*, Y_s^{1*})$ , depending on parameter values.

This ambiguity in the effects of better uniform enforcement on the diffusion of good values is a robust feature of the model and it is not just due to the unrestricted matching technology. In particular, suppose that all matches have the same probability irrespective of distance, so that  $g(y)$  is a constant for  $y \leq S$ . The appendix shows that the effect of  $q$  on educational effort remains ambiguous. Here is why. On the one hand, as shown in the appendix, better enforcement (lower  $q$ ) increases the difference in expected behavior of the two player types ( $Y_t^1 - Y^0$  rises). That is, at the margin better enforcement has a bigger effect on the behavior of the good than of the bad players. This induces more parental effort, because the consequences of having a kid with better values are more important. On the other hand, the improvement in the behavior of the bad player (the higher  $Y^0$ ) reduces the opportunity cost of having a bad kid, because the "warm glow" benefit of cooperation decreases with distance. In other words, the difference in behavior between a good and a bad kid, although more likely, occurs over more distant matches and thus it is less costly for the parent. Hence the incentives to put effort into education can either rise or fall as enforcement improves uniformly, depending on which of these two effect prevails.<sup>17</sup>

Summarizing, better external enforcement can either improve or destroy values, depending on where it bites. Distant transactions are mainly enforced by the law, since informal means are unlikely to be sufficiently powerful. If

---

<sup>17</sup>The appendix shows that, if  $Y_t^1 - Y^0$  is large, then the first effect prevails and  $f_t$  rises as enforcement improves. The opposite happens if  $Y_t^1 - Y^0$  is small.

so, an improvement in the quality of legal institutions is likely to crowd in better values. The reason is that better law enforcement makes moral behavior economically less costly, which encourages the diffusion of generalized morality.

Local transactions, instead, are likely to be enforced by informal methods, particularly when legal institutions are weak. But here an improvement of the local enforcement technology hurts the diffusion of generalized morality. The reason is that strong local enforcement makes it less necessary to rely on values to induce good behavior. This creates a substitutability between values and external enforcement, which weakens the incentive to diffuse generalized morality.

These results can explain why norms of limited morality are often observed in societies with a history of weak legal institutions, and with a recent legacy of semi-feudal arrangements. Southern Italy is a case in point. As described by Gambetta (1993), feudalism was formally abolished in 1806 in the continental South, and in 1812 in Sicily, much later than in the rest of Europe. This transformation and the associated introduction of private property rights for land created a demand for protection that the weak state institutions could not fulfill. The vacuum was soon exploited by the mafia in Sicily, and by a system of local patronage elsewhere (Davis 1975). Like the feudal structures that they replaced, these illegal or informal enforcement methods only operated locally. In fact, to increase the value of the protection it offered, the mafia drew a sharp line between those under its protection and everyone else, and made sure that those that refused its services would be easily abused (Gambetta 1993, chapter 1). In terms of the model, the mafia can be thought as better local enforcement (a reduction in  $q^0$ ) together with much worse enforcement of distant transactions (a higher  $q^1$ ). While such local enforcement methods can be effective in sustaining cooperation within the community, they are nevertheless very destructive of the values towards society at large. Indeed, several authors have documented the lack of generalized trust and the diffusion of values consistent with limited morality in Southern Italy (Banfield 1958, Putnam 1993, Guiso et al. 2008, Tabellini 2005, 2008).

### 3.4.2 Economic geography

Next, consider the effects of the matching technology, as captured by the probability of a match with someone located at distance  $y$ ,  $g(y)$ . With ref-



erence to Figure 1, this immediately shifts the curve  $N(Y_t^1, n_{t-1})$ , through the parents' incentives to educate their kid, while the curve  $Y(n_t)$  remains unaffected.

Specifically, consider a uniform increase in the probability of matches in the interval  $[Y^0, Y_s^1]$ , where the two types of players behave differently. Suppose that this occurs at the beginning of period 0 before parents choose effort. By (10), equilibrium effort  $f_0$  in period 0 increases. Intuitively, the interval  $[Y^0, Y_s^1]$  is where the difference between the two types is relevant, and hence where effort pays off from the parent's perspective. Hence, any increase in the probability of matches in this region induces more effort. This in turn increases the fraction of good players,  $n_0$ , which also brings about an immediate expansion of the upper threshold of cooperation,  $Y_0^1$ . From here onwards, convergence to the new steady state takes place, as described above, eventually leading to a higher fraction of good players,  $n_s$ , and a larger upper threshold of cooperation,  $Y_s^1$ .<sup>18</sup>

**The effects of localization** An increase in the probability of matches in the interval  $[Y^0, Y_s^1]$  can occur for different reasons, suggesting different interpretations. One possibility is that the probability of nearby matches, at distance below  $Y^0$ , drops. In other words, the economy has become less localized. Thus, the model suggests that norms of generalized morality become more diffuse if transactions are less localized, and viceversa that a more localized economy hurts the diffusion of good values.

Although the analytical mechanism is different, this result is clearly in line with the effects of local vs global enforcement discussed above: both local enforcement and local matching breed limited morality. In the model the matching technology is exogenous. In practice, however, the spatial patterns of economic activity and enforcement are likely to be jointly determined with self reinforcing effects. If enforcement is primarily local, this encourages local interactions, and viceversa, as shown by Dixit 2004, frequent local interactions facilitate informal enforcement methods that sustain local cooperation. The result that localization of enforcement and of transactions have the same effects on values is therefore a sign of robustness. Section 4 below illustrates yet another feedback effect, from values to enforcement.

Abandoning a literal interpretation, these results can contribute to ex-

---

<sup>18</sup>As can be seen by (10), a change in the matching technology  $g(y)$  that increases  $E[e^{-\theta^1 y} | Y_s^1 \geq y \geq Y^0]$  leaving  $\Pr(Y_s^1 \geq y \geq Y^0)$  unaffected has similar effects.

plain differences in traditional economic organizations between what Greif (1997) calls “collectivist” vs “individualist” societies. In the former, individuals interact primarily with members of their own group; enforcement is achieved through informal methods that only apply to group members; and values reflect group loyalty and neglect or intolerance towards members of other groups. In individualist societies, instead, group membership is not well defined as individuals mix and interact across groups; enforcement is achieved through formal and specialized means, such as courts, that can be used in a vast range of transactions; and the individual, rather than the group, is the primary repository of values and intrinsic rights. Greif (1994, 1997) contrasts the Maghribi and Genoese traders in the late medieval period as two examples of these different arrangements, where the spatial patterns of enforcement, transactions and values have mutually reinforcing effects in the organization of society.

Historical differences in the economic organization between East and West can be explained along similar lines. In Western Europe impersonal exchange took place in anonymous markets supported by specialized institutions obeying formal procedures. In East Asia markets were organized through a web of kin-based social structures linked by personal relations (Greif 2005). As suggested by the model, these different arrangements are likely to encourage the diffusion of different values: in the West a generalized respect for the individual and their rights, in the East a culture of loyalty to the local community or to a network of relatives and friends. In turn, as suggested by Greif (2005) and as modelled in section 4 below, these different values facilitated the evolution of different political and economic arrangements, with feedback effects in both directions.

Kumar and Matsusaka (2006) reason along similar lines but they go even further, arguing that the difference in economic geography between East and West (and the implied spatial pattern of economic transactions) is an exogenous force that brought about different economic organizations. They quote historical evidence that in the XVIIth century population density was much higher in China and India than in Europe. Moreover, long distance travel was easier within Europe than within the far East, because of both geography and relative availability of means of transportation. As a result, trade was more frequently local in Asia than in Europe. As suggested by the model of this paper, this might explain the greater diffusion of norms of generalized morality in the West than in the East.

**The effects of globalization** Alternatively, the probability of matches in the region  $[Y^0, Y_s^1]$  can go up because very distant matches (above  $Y_s^1$ ) have become less likely. This too is beneficial to the diffusion of good values, because the more frequent interactions inside the community of reference for the good players strengthens their incentive to transmit these values to their offspring.

Taken literally, this says that globalization (the equivalent of more frequent very distant matches) reduces the scope of cooperation, because it destroys the values that induce individuals to cooperate. It should be acknowledged that here the model formulation is particularly constraining, however. By assumption, there are only two types of players, and parents can only mould the parameter  $\theta$  in the model. This constrains how the diffusion of values might be affected by changes in the matching technology.

More generally, this result can be interpreted as saying that the diffusion of good values is hurt by economic forces that induce individuals to move outside of the community with which older generations identify. This is consistent with observations by economic historians like Polanyi (1957), that the industrial revolution destroyed moral values in the UK because it brought about large social dislocations. More recently, Miguel et al. (2002) have studied the effect of industrialization in Indonesia since the mid 1980s. Exploiting repeated cross sections of national surveys, they contrasted indicators of social capital, trust and community values in districts that experienced rapid industrialization and in neighboring districts. The districts that were left behind and with severe out-migration suffered a large destruction of social capital and community values. This is what the model would predict, although, if taken literally, over a longer time span than observed in Indonesia.

Summarizing, the general insight of the model is that the evolution of values reflects the patterns of economic interactions relative to the pattern of moral ties between individuals. Whatever increases the likelihood of interactions in the region between  $Y^0$  and  $Y_t^1$ , where the distinction between limited and generalized morality matters, also increases the diffusion of trustworthiness within the community. Very local interactions (below  $Y^0$ ) or very distant interactions (above  $Y_t^1$ ) have the opposite effect, because the distinction between limited and generalized morality has no behavioral implication in those regions, and this dampens the incentive to invest in good values.

If the matching technology varies systematically with the stages of development, this suggests a non-monotonic relationship between economic

development and values (or the scope of cooperation). At early stages of development, transactions are mainly local, and both values and cooperation remain more limited in scope. As development progresses, and impersonal transactions gain relevance, this is accompanied by a generalization of the scope of values and cooperation. But at even more advanced stages of development, transactions become so spread out relative to community values that the negative effects of globalization on values and cooperation might gain relevance.

More generally, over time values adapt to the spatial pattern of economic incentives, induced by either external enforcement or the likely economic transactions. Limited morality prevails when individuals interact locally, and local transactions are better enforced than anonymous market exchange. Viceversa, generalized respect for others flourishes if enforcement is effective also outside the local community, and interactions are more widespread (but not too spread out relative to the domain of current values in the community).

## 4 Endogenous Enforcement

The previous section highlighted how external enforcement in local vs distant matches affects the transmission of values. This section focuses on the opposite feedback effects: how the distribution of values influences the spatial patterns of external enforcement. As argued above, local enforcement is likely to be achieved through informal means, whereas distant enforcement reflects the functioning of state institutions. Thus, I view the quality of distant enforcement as determined by the political process. For simplicity I take local enforcement as given, and I consider a simple model of majoritarian politics, where citizens vote on the quality of distant enforcement.

The central result is that politics and culture interact with mutually reinforcing effects. The previous section argued that weak state institutions (i.e. poor enforcement of distant transactions) induce an adverse distribution of values. Here I show that initially adverse values lead to weak state enforcement. In this explicitly dynamic economy, not only there can be multiple equilibria, but there is also hysteresis: initial conditions matter, because they lead the economy to different steady states.<sup>19</sup>

---

<sup>19</sup>A similar point is illustrated with respect to welfare state policies in models by Bisin and Verdier (2000, 2004), where individual tastes for private vs public consumption or leisure vs work are endogenous, and by Bénabou and Tirole (2006a), where parents conceal

## 4.1 The enforcement regime in political equilibrium

To simplify the algebra, I make two assumptions. First, the probability of a match at distance  $y$  is the same for any distance:  $g(y) = \eta$ . Second, the material payoffs in the prisoner's dilemma game satisfy  $l = w$ . Thus, by Proposition 2 the matching game has a unique equilibrium. In terms of Figure 1, the  $Y(n_t)$  curve is vertical. This also means that any strategic complementarity can only arise from the endogeneity of government policy, since for a given policy the equilibrium is unique.

I model external enforcement as in the previous section, namely as a probability  $1 - q(y)$  that cheating is detected, where detections erase the gain from cheating and the loss from being cheated. For matches up to a right hand neighborhood of  $Y^0$ ,  $q(y) = q^0$ , with  $q^0$  a given parameter reflecting informal local arrangements. For more distant matches,  $q(y) = q_t^1$ , and  $q_t^1$  is determined by government policy subject to  $q^0 \leq q_t^1 \leq Q$ , where  $Q > q^0$  is a fixed parameter. Since a smaller  $q$  corresponds to better enforcement, state enforcement of distant transactions can be worse than local informal arrangements (if  $q_t^1 > q^0$ ), but not strictly better.

Given this formulation and repeating the analysis in section 2, the maximum distance that sustains cooperation by the good and bad players respectively is given by:

$$Y_t^k = \lceil \ln(d/wq_t^k) \rceil / \theta^k, \quad k = 0, 1 \quad (17)$$

Government policy is set under majority rule in each period. The timing of events is as follows. First, parents choose effort. Then, the kids' type becomes known and the kids vote over the enforcement regime (parents don't participate in the vote). Finally, the kids play the matching game. Note that, under this timing, the kids only consider their utility in the current period. When the vote is taken, the fraction of good players is already determined. Thus, current enforcement only affects current expected payoffs.

What is the policy preferred by the two types? It is easy to verify that good players always prefer the strongest possible state enforcement, corresponding to  $q_t^1 = q^0$ , since this reduces their loss from being cheated (see subsection 6 of the Appendix). We call this outcome the *strong enforcement regime*.

Bad players, instead, face a tradeoff: on the one hand, worse state enforcement (a higher  $q_t^1$ ) increases the benefit of cheating; on the other hand, 

---

 information to their kids to overcome a time-inconsistency problem.

it makes the good players more cautious, and this in turn shrinks the range of matches over which the bad players can take advantage of a cooperating opponent. Subsection 6 of the Appendix proves that the first effect dominates at the lower bound  $q^0$ , if the psychological cost  $d$  is sufficiently large relative to the material payoffs of the prisoner's dilemma game (see condition A2 in the appendix). Hence, under this condition, bad players always prefer a weaker state enforcement than technologically feasible,  $q_t^1 > q^0$ . The appendix also shows that the optimal policy from the point of view of the bad players is time invariant, since it does not depend on  $n_t$ . We call this policy outcome the *weak enforcement regime*, and we denote it as  $q_t^1 = \bar{q}$ , where  $\bar{q} > q^0$ .<sup>20</sup>

Given these results, the political equilibrium in any period  $t$  is straightforward and it is summarized in the following:

**Lemma 6** *Suppose that condition (A2) in the appendix holds. If  $n_t > 1/2$  then the strong enforcement regime prevails in period  $t$ :  $q_t^1 = q^0$ . If  $n_t < 1/2$  then the weak enforcement regime prevails in period  $t$ :  $q_t^1 = \bar{q}$ , with  $\bar{q} > q^0$ . If  $n_t = 1/2$  then either regime can prevail.*

## 4.2 Equilibrium dynamics

As discussed in the previous section, the regime with weak state enforcement dampens the incentives to inculcate trustworthiness. Specifically, let  $f$  and  $\bar{f}$  denote effort under strong and weak enforcement respectively (since with  $w = l$  effort no longer depends on  $n_t$ , time indexes are removed). Subsection 7 of the appendix proves that parents exert more effort in the strong than in the weak enforcement regime:

**Lemma 7**  $\bar{f} = f - \Delta > 0$  with  $\Delta > 0$

This set up induces a strategic complementarity in the education decision of the parents. If parents expect  $n_t > 1/2$ , then they anticipate better enforcement and, by Lemma 7, they exert more effort to inculcate trustworthiness. This in turn increases the fraction of good players, and might bring

---

<sup>20</sup>For simplicity I have assumed that state enforcement comes for free. It would be straightforward to add a cost to improve state enforcement; if the cost was shared equally by all players, this would weaken condition A2, since now the bad players would have an additional reason to prefer weak state enforcement.

about a political equilibrium where they are a majority. Viceversa, if parents expect  $n_t < 1/2$ , they reduce effort, which might shift future political majorities. For some parameter values, this can give rise to multiple steady states.

Specifically, suppose that parents expect strong enforcement. Then the steady state fraction of good players is given by (16) in the previous section, reproduced here for convenience (with  $F(Y_s^1)$  replaced by  $f$ ):

$$n_s^* = \frac{\delta}{1-f} \quad (18)$$

If  $n_s^* > 1/2$ , this steady state reproduces itself in a political equilibrium. Suppose instead that parents expect weak enforcement. Then, by Lemma 7, the steady state fraction of good players is:

$$\bar{n}_s^* = \frac{\delta}{1-f+\Delta} \quad (19)$$

If  $\bar{n}_s^* < 1/2$ , this steady state too reproduces itself in a political equilibrium. Thus, both steady states are possible in equilibrium if  $n_s^* > 1/2 > \bar{n}_s^*$ , or, by (19) and (18), if:

$$\Delta > 2\delta + f - 1 > 0 \quad (A3)$$

Note that, since  $f > \Delta$ , the left hand inequality requires  $\delta < 1/2$ .

As already noted, if  $w = l$ , then the curve  $Y(n_t)$  in Figure 1 is vertical, and neither  $Y^1$  nor effort depend on  $n_t$ . Thus both steady states are dynamically stable and the adjustment is monotonic. Which steady state is reached in equilibrium depends on the initial conditions and on parents' expectations, as we now discuss.

Strong enforcement is a political equilibrium in period  $t$  if, given that it is expected, we have  $n_t > 1/2$ . By (8), this condition can be stated as:

$$n_t = \delta + n_{t-1} f > 1/2 \quad (20)$$

Similarly, weak enforcement is a political equilibrium in period  $t$  if, given that it is expected,  $n_t < 1/2$ , namely if :

$$n_t = \delta + n_{t-1}(f - \Delta) < 1/2 \quad (21)$$

Combining (20) and (21), we obtain two thresholds, that define which equilibria exist in period  $t$ , depending on the fraction of good players in period  $t - 1$ . Specifically, let:

$$\hat{n} = \frac{1 - 2\delta}{2f} \quad (22)$$

$$\hat{N} = \frac{1 - 2\delta}{2(f - \Delta)} \quad (23)$$

with  $\hat{N} > \hat{n}$ . Then we have:

**Lemma 8** *If  $n_{t-1} < \hat{n}$ , then in period  $t$  the unique equilibrium has weak enforcement. If  $n_{t-1} > \hat{N}$  then in period  $t$  the unique equilibrium has strong enforcement. If  $\hat{N} \geq n_{t-1} \geq \hat{n}$ , then both the weak and the strong enforcement regimes exist as equilibria in period  $t$ .*

The proof is straightforward. If  $n_{t-1}$  is so low that it falls below the threshold  $\hat{n}$ , then even if parents expect strong enforcement, we have  $n_t < 1/2$ . Hence strong enforcement cannot be a political equilibrium. Conversely, if  $n_{t-1}$  is so high that it exceeds the threshold  $\hat{N}$ , then even if parents expect weak enforcement we have  $n_t > 1/2$ , which rules out weak enforcement as an equilibrium. For values of  $n_{t-1}$  in between the two thresholds, either regime could win a majority depending on the parents' expectations.

Suppose that condition (A3) is satisfied, so that we have two steady states. Suppose further that both steady states fall outside of the interval  $[\hat{n}, \hat{N}]$ . Manipulating (22)-(23) and (18)-(19), a sufficient condition for this to happen is:

$$1 - \frac{1-f}{2\delta} > \Delta > \frac{f}{1-2\delta} - 1 \quad (A4)$$

which in turn requires  $\delta \leq 1/4$  (and which also implies (A3)). Since the adjustment towards the steady state is monotonic, then the thresholds  $\hat{n}$  and  $\hat{N}$  define three regions with different dynamics. If the economy starts from an initial condition  $n_0 < \hat{n}$ , then the equilibrium is unique. The economy remains forever in the weak enforcement equilibrium and it converges to the weak enforcement steady state. Conversely, if the economy starts from an initial condition  $n_0 > \hat{N}$ , then the equilibrium is again unique. The economy remains forever in the strong enforcement equilibrium and it converges to the strong enforcement steady state. If the initial condition is in between these two thresholds,  $n_0 \in [\hat{n}, \hat{N}]$ , then both paths are feasible equilibria, and the economy eventually ends up in one or the other steady state depending on expectations.



If condition (A4) is violated, then one of the steady states (or both) are inside the region where multiple equilibria are possible. In this case eventually the economy might end up in the region of multiple equilibria, and one or the other steady state will be reached depending on expectations (if both inequalities in (A4) are violated then both steady states are inside the region of multiple equilibria and the economy certainly reaches this region in finite time for any initial conditions).

We summarize the foregoing discussion in the following.

**Proposition 9** *If condition (A3) holds, then the economy has two steady states, one with strong external enforcement and where the good players are a majority; and one with weak external enforcement and where the good players are a minority. Both steady states are dynamically stable. If condition (A4) also holds, and if the initial fraction of good players,  $n_0$ , is outside of the interval  $[\hat{n}, \hat{N}]$ , then the equilibrium is unique. For  $n_0 < \hat{n}$  (for  $n_0 > \hat{N}$ ), the economy remains always under the weak (strong) enforcement regime and eventually reaches the weak (strong) enforcement steady state. If condition (A4) is violated, then multiple equilibria are possible during the adjustment path towards one or the other steady states.*

### 4.3 Discussion

Proposition 9 highlights the importance of mutually reinforcing effects between culture and politics when both are endogenous. On the one hand, effective law enforcement strengthens the incentives to transmit sound values. On the other hand, the quality of law enforcement is also endogenous, and reflects deliberate policy choices. A society with weak values, or where respect for the law and for others is lacking, is also more tolerant of lax law enforcement. As a result, otherwise identical societies may end up along very different paths if they start from different initial conditions.

Thus, this Proposition can explain why distant historical circumstances have such long lasting effects, and why some societies may remain trapped in cultural, institutional and economic backwardness. Weak state enforcement forces citizens to rely on informal arrangements that sustain local but not global cooperation. This diffuses adverse cultural traits in the community, which in turn influence political outcomes. Weak state enforcement is retained even under democracy, because adverse cultural traits make citizens more tolerant of ineffective government. Better enforcement institutions are

available, and nothing prevents citizens from adopting them, but this does not happen in a political equilibrium. Whether lax law enforcement refers to cheating on a private transaction, or tax evasion, or free riding on a public good, it is tolerated and perhaps even preferred by a majority of citizens. This cultural explanation of institutional persistence is quite different from others suggested in the literature, that emphasize the power of the *élites* against the will of the citizens at large (eg. Acemoglu and Robinson 2006).

The presence of at least some citizens who value cooperation and who are occasionally exploited by other more shrewed players is not necessary for this result. Even if almost everyone ends up with a low value for cooperation, better law enforcement would still be opposed if it costs resources. The reason is that the benefits of better enforcement would be negligible in a society where trust and cooperation are so low that many mutually advantageous trade opportunities are foregone anyway. A similar result would also hold if political outcomes reflect the influence of organized interests, such as the mafia or local patrons, who extract rents from providing local protection, and hence benefit from weak state enforcement. Thus, majoritarian politics is not the only way to generate the feedback effects discussed here.

Strictly speaking, this model focuses on cooperation in a bilateral transaction. But a similar tradeoff between values and incentives is also at work in multilateral cooperation or generalized public good provision. A particularly relevant example is the act of voting to discipline a political agency problem. Ousting a corrupt politician, or selecting a competent leader, requires citizens to be informed, to bear the cost of voting, and to vote according to general social welfare rather than their own particularistic benefit. As shown for instance by Ferejohn (1986), this can only be achieved through implicit cooperation amongst the voters. In particular, each individual voter or group must refrain from rewarding a corrupt politician who offered targeted personal benefits in exchange for votes. These forms of political cooperation may not be sustainable where limited morality prevails. If government abuse and nepotism in turn induce the diffusion of adverse cultural traits, then we have yet another loop where political and cultural outcomes have mutually reinforcing effects. Although this remains to be modeled more precisely, preliminary evidence in Tabellini (2008) supports this idea. Italian voters in regions where generalized morality is more diffused, and that were ruled by better political institutions over two centuries ago, today are more willing to punish incumbent politicians under criminal investigation.

The interaction between culture and politics has relevant implications also

for group formation and redistributive policies. It is well known, for instance, that in Africa public policies often provide targeted benefits to ethnic groups. Opportunistic politicians have an incentive to do so if individuals identify with ethnic groups, rather than with groups formed along other economic or social dimensions. But group identity is not exogenous, on the contrary, it is likely to be strengthened by any policy that targets the group. Indeed, Miguel and Posner (2006) found that in Africa ethnic identity is stronger amongst those who are more likely to be exposed to public policies. The approach of this paper could be extended to examine the historical reasons that make some groups influential, and to study the joint evolution of redistributive politics and group identity.

Finally, in the model individuals were assumed to vote according to their self interest. Hence cultural traits influence political preferences only through economic behavior, because this determines how individuals are affected by external enforcement. If instead individual values also have a direct impact on political ideologies, as seems plausible, then there is an additional channel through which values might shape public policies, again with feedback effects in both directions.<sup>21</sup>

## 5 Concluding remarks

I conclude by discussing several possible extensions of this basic framework, and other recent related work.

The model literally assumes that values are transmitted within the family, and that only parents make purposeful educational choices. In practice, other channels of cultural transmission, from peers, own experience, educational institutions or the media, are also likely to be important. This opens the door to other relevant choices, such as whom to select as your friend, or how intensely to experiment. It also gives a role to other motivated actors who might have economic or political reasons to influence cultural traits. Recent work by Anderlini et al. (2007), Bénabou (2008), Fernandez (2007b) and Guiso et al. (2008) has started investigating the role of belief formation and manipulation in a variety of related settings.

---

<sup>21</sup>Alesina and Angeletos (2005) consider a model where individuals vote according to their self interest and also according to a notion of what is fair and unfair. In their model, however, individual values are exogenously given and do not interact with the economic environment.

I have also neglected reputational forces, not because they are unimportant, but to focus on values alone. As stressed for instance by Kaplow and Shavell (2007), values also interact with reputational incentives. Sustained punishment of deviant behavior is more likely to be incentive compatible, or to be a focal point of coordinated action, if the deviation is morally tainted. Thus, reputational mechanisms support law enforcement when the law is considered fair, or if the state enjoys the confidence of citizens, but not if individual values clash with the state. Incorporating these channels in this model, perhaps exploiting the work of Dixit (2004), is feasible and might lead to new insights.

Reputation can also operate through signalling. If values were observable, then players with good values might have an advantage in this model, because they could induce their partner to cooperate over a larger range of matches. This would change the incentives of parents, and even the bad players in the model might want to transmit better values to their children, depending on the strength of external incentives. True values may be unobservable, but individuals may find ways of signalling them (other than through repeated cooperation). Levy and Razin (2006), for instance, formulate a theory of religion based on the assumption that religious rituals are observable (maybe only within a subset of the population). This creates a strategic incentive to join a religious organization, to signal one's type. In Levy and Razin (2006), individual values are stable and exogenous and individuals choose their own religion (i.e. there is no role for parents to shape their kid's values or beliefs). Combining their insights with the model of this paper seems doable and promising.<sup>22</sup>

In summary, much remains to be done to pin down more precisely the channels of cultural transmission both inside and outside of the family, to understand the role of learning and formal education, and to study empirically the implications of specific cultural traits. These issues can be fruitfully studied with the standard tools of economic analysis, and can yield important new insights on why cooperation is easier to sustain in some social environments than in others.

---

<sup>22</sup>Along similar lines, Baron (2007) has studied the role of a non-profit agency that certifies the players' types, in a version of this model with exogenous values.

## 6 Appendix

### 6.1 Proof of Lemma 2

By Proposition 1, the kid's expected material payoffs in the equilibrium of the matching game are:

$$U_t^k = \int_0^{Y_t^k} g(z)[c\pi_t(z) + (h-l)(1-\pi_t(z))]dz + \int_{Y_t^k}^S g(z)[(c+w)\pi_t(z) + h(1-\pi_t(z))]dz \quad (24)$$

where  $\pi_t(z)$  denotes the probability that a partner at distance  $z$  will cooperate in period  $t$  in the Pareto superior equilibrium -  $\pi_t(z)$  is indexed by time because it might depend on  $n_t$ . The first term on the RHS is the expected utility when cooperating, given that the partner cooperates with probability  $\pi_t(z)$ . The second term is the expected utility of not cooperating, again given the probability that the partner cooperates.

Consider the solution to the problem of maximizing  $V^{pk}$ , as defined in (7), by choice of  $\theta^k$  - we now omit time indexes as they are redundant. As discussed in the text,  $\theta^k$  enters the expression for  $V^{pk}$  only through the distance threshold  $Y^k$  that triggers non-cooperation by the kid. Hence, by (24) and (7), differentiating  $V^{pk}$  with respect to  $\theta^k$  and rearranging, we have:

$$\frac{\partial V^{pk}}{\partial \theta^k} = g(Y^k) \frac{\partial Y^k}{\partial \theta^k} \left\{ de^{-\theta^p Y^k} - [(w-l)\pi(Y^k) + l] \right\} \quad (25)$$

By (4)-(6),  $\frac{\partial Y^k}{\partial \theta^k} \leq 0$ , with strict inequality if  $Y^1 > Y^0$  (i.e. under A0). Hence, if  $Y^1 > Y^0$ , then the optimal value of  $d^k$  is such that  $Y^k$  solves the expression

$$de^{-\theta^p Y^k} = [(w-l)\pi(Y^k) + l] \quad (26)$$

for  $\pi(Y^k)$  corresponding to the equilibrium probability of cooperation by a partner located at distance  $Y^k$ . But by (3), this implies  $\theta^k = \theta^p$ . Hence the parent strictly prefers to have a kid with his/her own value parameter. If instead  $Y^1 = Y^0$ , then the parent is indifferent between a good or a bad kid. QED

## 6.2 Derivation of (10) and proof of Lemma 4

In equilibrium, (24) for  $k = 0, 1$  can be rewritten as:

$$U_t^0 = c \int_0^{Y^0} g(z) dz + [(c+w)n_t + h(1-n_t)] \int_{Y^0}^{Y_t^1} g(z) dz + h \int_{Y_t^1}^S g(z) dz \quad (27)$$

$$U_t^1 = c \int_0^{Y^0} g(z) dz + [cn_t + (h-l)(1-n_t)] \int_{Y^0}^{Y_t^1} g(z) dz + h \int_{Y_t^1}^S g(z) dz \quad (28)$$

The first term on the RHS of (27) and (28) corresponds to the expected material benefits in a match where both partners always cooperate; the second term is the expected outcome in the intermediate area where only the good players cooperate; the third term is the expected outcome where no player cooperates.

(27) and (28) imply:

$$\begin{aligned} U_t^1 - U_t^0 &= -[l + (w-l)n_t] \Pr(Y_t^1 \geq z \geq Y^0) = \\ &= -de^{-\theta^1 Y_t^1} \Pr(Y_t^1 \geq z \geq Y^0) < 0 \end{aligned} \quad (29)$$

where the last equality follows from (4)-(6) when  $Y^1 > Y^0$ . Inserting this in (9) yields (10).

Finally, to prove Lemma 4, differentiate the RHS of (10) with respect to  $Y_t^1$ . Simplifying, we have:

$$F_{Y_t^1} = \varphi d\theta^1 e^{-\theta^1 Y_t^1} \int_{Y^0}^{Y_t^1} g(z) dz > 0 \quad (30)$$

## 6.3 Slope of the functions $n_t = N(Y_t^1, n_{t-1})$ and $Y_t^1 = Y(n_t)$

Equation (12) implies that  $N_{Y_t^1} = n_{t-1} F_{Y_t^1} > 0$

Differentiating the RHS of (11) with respect to  $n_t$ , we also have:

$$Y_{n_t} = \frac{1}{\theta^1} \frac{l-w}{x_t} \geq 0 \quad (31)$$

where  $x_t = l + (w - l)n_t \geq 0$  (since  $n_t \leq 1$ ). Thus, the sign of  $Y_{n_t}$  is the same as that of  $l - w$ .

The function  $N(Y_t^1)$  intersects the function  $Y(n_t)$  from left to right, as drawn in Figure 1, if  $N_{Y_t^1} < 1/Y_{n_t}$ , or, by (26), (30) and (31), if:

$$\frac{1}{l - w} > \varphi n_{t-1} \int_{Y^0}^{Y_t^1} g(z) dz \quad (32)$$

Since  $n_{t-1} \int_{Y^0}^{Y_t^1} g(z) dz < 1$ , a sufficient for (32) to hold is (A1).

## 6.4 Dynamic stability of the steady state

Applying the implicit function theorem to (13) and (14), we have:

$$\begin{aligned} \frac{dn_t^*}{dn_{t-1}} &= \frac{N_{n_{t-1}}}{1 - Y_{n_t} \cdot N_{Y_t^1}} \\ \frac{dY_t^{1*}}{dn_{t-1}} &= \frac{dn_t^*}{dn_{t-1}} Y_{n_t} \end{aligned}$$

Under (A1),  $1 - Y_{n_t} \cdot N_{Y_t^1} > 0$ . Moreover:

$$N_{n_{t-1}} = f_t > 0$$

where the inequality follows from Corollary 3. Hence, (A1) implies  $\frac{dn_t^*}{dn_{t-1}} > 0$ . Moreover, since  $\text{sign}[Y_{n_t}] = \text{sign}[l - w]$ , we also have  $\frac{dY_t^{1*}}{dn_{t-1}} \geq 0$ , with strict inequality if  $l > 0$ .

To prove that  $\frac{dn_t^*}{dn_{t-1}} < 1$ , we need to prove that  $N_{n_{t-1}} < 1 - Y_{n_t} \cdot N_{Y_t^1}$ , or equivalently, that

$$f_t + Y_{n_t} \cdot N_{Y_t^1} < 1 \quad (33)$$

By (30) and (31), the term  $Y_{n_t} \cdot N_{Y_t^1}$  is proportional to  $\varphi$ . By (10), the term  $f_t$  is also proportional to  $\varphi$ . Define  $\bar{\varphi}_t$  as the value of  $\varphi$  such that  $f_t + Y_{n_t} \cdot N_{Y_t^1} = 1$ . Note that  $\bar{\varphi}_t$  depends on  $t$  through the terms  $x_t$  and  $n_{t-1}$  (cf. (31), (30) and (10)). Define  $\bar{\varphi} = \arg \min(\bar{\varphi}_t)$ , where the minimization is taken over all feasible values of  $n_{t-1}$  and  $x_t$ . Since  $N_{n_{t-1}} > 0$ ,  $Y_{n_t} \geq 0$  and  $N_{Y_t^1} > 0$ , then  $\bar{\varphi} > 0$ . Then, for any  $0 < \varphi < \bar{\varphi}$ , (33) also holds. QED

## 6.5 Enforcement and endogenous values

Let enforcement be as described in the text, subsection 3.4. Repeating the steps in section 2, we can rewrite the distance threshold  $\tilde{y}^k$  that leaves a player of type  $k$  just indifferent between cooperating or not as:

$$\tilde{y}^k = \{ \ln d - \ln[(w-l)\pi(\tilde{y}^k) + l] q^k \} / \theta^k \quad (34)$$

**Local vs distant enforcement** Consider the case in  $q^0$  and  $q^1$  vary independently of each other in a neighborhood of  $q^0 = q^1$ . Replace  $w$  with  $wq^k$  and  $l$  with  $lq^k$  respectively in (4), (5), (6) and then differentiate with respect to  $q^k$ . We have:

$$\frac{\partial Y^k}{\partial q^k} = -1/q^k \theta^k, \quad \text{for } k = 0, 1 \text{ and } \frac{\partial Y^k}{\partial q^h} = 0 \text{ for } k \neq h \quad (35)$$

By (10), it can then easily be verified that:  $\frac{\partial f_t}{\partial q^0} > 0 > \frac{\partial f_t}{\partial q^1}$  (holding  $n_t$  constant). The discussion in the text then follows.

**Uniform enforcement** Consider now the case  $q^0 = q^1 = q$ , and consider the effect of a change in  $q$ . The result in (35) now becomes:

$$\frac{\partial Y^k}{\partial q} = -1/q \theta^k, \quad \text{for } k = 0, 1 \quad (36)$$

Since  $\theta^0 > \theta^1$ , this implies  $\frac{\partial Y^1}{\partial q} < \frac{\partial Y^0}{\partial q} < 0$ . In other words, at the margin  $Y^1$  is more reactive than  $Y^0$  to changes in  $q$ .

Next, differentiate the right hand side of (10) with respect to  $q$ , holding  $n_t$  fixed and exploiting (36). After some simplifications we obtain:

$$\frac{\partial f_t}{\partial q} = \frac{\varphi d}{q \theta^0} e^{-\theta^1 Y_t^1} \left\{ g(Y^0) \left[ e^{\theta^1(Y_t^1 - Y^0)} - 1 \right] - \theta^0 \int_{Y^0}^{Y_t^1} g(z) dz \right\} \quad (37)$$

The sign of (37) is generally ambiguous. The ambiguity remains even if we assume that the density  $g(y)$  does not depend on  $y$  (i.e. that the probability of a match at distance  $y$  is uniform).

Specifically, let  $g(y) = \eta$ . Then simplifying (37) we have:

$$\text{sign} \left\{ \frac{\partial f_t}{\partial q} \right\} = \text{sign} \left\{ e^{\theta^1 a_t} - 1 - \theta^0 a_t \right\} \quad (38)$$



where  $a_t = Y_t^1 - Y^0$ . At the point  $a_t = 0$  the RHS of (38) is zero. The derivative of the RHS of (38) with respect to  $a_t$  in a neighborhood of  $a_t = 0$  equals  $\theta^1 - \theta^0$ , which is strictly negative as  $\theta^0 > \theta^1$ . Hence, for  $a_t > 0$  but small,  $\frac{\partial f_t}{\partial q} < 0$ . But as  $a_t$  keeps rising eventually  $\frac{\partial f_t}{\partial q} > 0$ .

## 6.6 Proof of Lemma 6

Here I show that the good players always prefer the strong enforcement regime,  $q_t^1 = q^0$ , and I provide a sufficient condition that guarantees that the bad players prefer the weak enforcement regime,  $q_t^1 = \bar{q} > q^0$ .

Let  $U_t^k$  denote the expected utility of players of type  $k$  as a function of  $q_t^1$ . Adapting (28) and (7) to the new notation, we have:

$$\begin{aligned} \frac{\partial U_t^1}{\partial q_t^1} &= -w(1 - n_t) \int_{Y^0}^{Y_t^1} g(z) dz + \\ &+ g(Y_t^1)(c - h + wq_t^1)n_t \frac{\partial Y_t^1}{\partial q_t^1} < 0 \end{aligned}$$

where we have used  $wq_t^1 = de^{-\theta^1 Y_t^1}$  and where the negative sign follows from  $\frac{\partial Y_t^1}{\partial q_t^1} = -1/\theta^1 q_t^1 < 0$ . Since this expression holds for any  $q_t^1 \geq q^0$ , the good players are always in favor of the lowest possible value of  $q_t^1$ , namely they prefer the weak enforcement regime  $q_t^1 = q^0$ .

Next, consider the bad players. Since the threshold  $Y^0$  is not affected by the regime, the derivative of their expected utility with respect to  $q_t^1$  is:

$$\frac{\partial U_t^0}{\partial q_t^1} = wn_t \int_{Y^0}^{Y_t^1} g(z) dz + (c - h + wq_t^1)n_t \frac{\partial Y_t^1}{\partial q_t^1} g(Y_t^1) \quad (39)$$

We assumed  $g(z) = \eta$ . Moreover, by (17),  $\frac{\partial Y_t^1}{\partial q_t^1} = -1/\theta^1 q_t^1$ . Hence, (39) simplifies to:

$$\frac{\partial U_t^0}{\partial w_t} = wn_t \eta (Y_t^1 - Y^0) - n_t \eta \frac{(c - h + wq_t^1)}{\theta^1 q_t^1} \quad (40)$$

Evaluating the right hand side of (40) at the point  $q_t^1 = q^0$ , and simplifying, we can show that it is strictly positive if :

$$\ln d > \ln wq^0 + \frac{(c - h + wq^0)\theta^0}{wq^0(\theta^0 - \theta^1)} \quad (\text{A2})$$

By (17),  $Y_t^1$  only depends on time through  $q_t^1$ . Hence, by (40), the optimal value of  $q_t^1$  from the point of view of the bad players is constant. Denoting such optimal value by  $\bar{q}$ , under (A2) we have  $\bar{q} > q^0$ . QED

## 6.7 Proof of Lemma 7

Let an upper bar over a variable denote the corresponding variable in the weak enforcement regime. Repeating the previous analysis, parental effort in the weak enforcement regime is given by the following first order condition, adapted from (9):

$$\bar{f} / \varphi = (\bar{U}^1 - \bar{U}^0) + d \int_{Y^0}^{\bar{Y}^1} e^{-\theta^1 z} g(z) dz \quad (41)$$

where time subscripts have been dropped because under the simplifying assumptions of this section  $n_t$  no longer enters the right hand side of (41). We know from previous results that  $\bar{f} > 0$ . Moreover, in the weak enforcement regime, the difference in the expected material payoffs of a good and a bad kid can be written as (cf. subsection 2 of the appendix):

$$\begin{aligned} \bar{U}^1 - \bar{U}^0 &= -w\bar{q} \int_{Y^0}^{\bar{Y}^1} g(z) dz = \\ &= -wq^0 \int_{Y^0}^{Y^1} g(z) dz + w(q^0 - \bar{q}) \int_{Y^0}^{Y^1} g(z) dz + w\bar{q} \int_{\bar{Y}^1}^{Y^1} g(z) dz \quad (42) \end{aligned}$$

Combining (42) and (41), and exploiting (10) for  $w = l$ , we have:

$$\bar{f} = f - \Delta$$

where, under the simplifying assumption that  $g(z) = \eta$  :

$$\Delta = \varphi\eta \left\{ w(\bar{q} - q^0)(Y^1 - Y^0) - w\bar{q}(Y^1 - \bar{Y}^1) + d \int_{\bar{Y}^1}^{Y^1} e^{-\theta^1 z} dz \right\} \quad (43)$$

We also have:

$$\begin{aligned} \frac{\partial \Delta}{\partial \bar{q}} &= \varphi\eta \left\{ w(\bar{Y}^1 - Y^0) + w\bar{q} \frac{\partial \bar{Y}^1}{\partial \bar{q}} - d e^{-\theta^1 \bar{Y}^1} \frac{\partial \bar{Y}^1}{\partial \bar{q}} \right\} = \\ &= \varphi\eta w(\bar{Y}^1 - Y^0) > 0 \end{aligned}$$

where the second equality follows from  $d = w\bar{q} e^{\theta^1 \bar{Y}^1}$ , and where  $(\bar{Y}^1 - Y^0) > 0$  follows from the optimality condition (40) and the definition of  $\bar{q}$  as the optimal value of  $q_t^1$  for the bad players. Note that when  $\bar{q} = q^0$ , we have  $\bar{Y}^1 = Y^1$  so that  $\Delta = 0$ , by (43). Thus,  $\bar{q} > q^0$  implies  $\Delta > 0$ . QED

## References

- [1] Acemoglu, Daron, Simon Johnson, and James A. Robinson, "The Colonial Origins of Comparative Development : An Empirical Investigation" *American Economic Review*, 91 (2001), 1369-1401.
- [2] Acemoglu, Daron, and James A. Robinson, "Persistence of Power, Elites and Institutions", NBER Working Paper No. 12108, March 2006.
- [3] Akerlof, George A., and Rachel E. Kranton, "Economics and Identity," *Quarterly Journal of Economics*, 115 (2000), 715-733.
- [4] Akerlof, George A., and Rachel E. Kranton, "*Economics and Identity*", Mimeo, (2006), University of California, Berkeley.
- [5] Alesina, Alberto, and Eliana La Ferrara, "Participation in Heterogeneous Communities", *The Quarterly Journal of Economics*, 115 (2000), 847-904.
- [6] Alesina, Alberto, and Eliana La Ferrara, "Who Trusts Others?", *Journal of Public Economics*, 85 (2002) , 207-234.
- [7] Alesina, Alberto, and George-Marios Angeletos, "Fairness and Redistribution", *American Economic Review*, 95 (2005), 960-980.

- [8] Algan, Yann, and Pierre Cahuc, "Why Is The Minimum Wage So High In Low-Trust Countries?", Mimeo, (2006), University of Paris 1.
- [9] Algan, Yann, and Pierre Cahuc, "Social Attitudes and Macroeconomic Performance: an Epidemiological Approach", Mimeo, (2007), University of Paris 1
- [10] Amir, Rabah, "Supermodularity and Complementarity in Economics: An Elementary Survey", Mimeo, (2003), CORE.
- [11] Anderlini, Luca, Dino Gerardi and Roger Lagunoff, "Social Memory and Evidence from the Past," mimeo, (2007), Georgetown University.
- [12] Andreoni, James, "Impure Altruism and Donations of Public Goods: A Theory of Warm-Glow Giving", *The Economic Journal*, 100 (1990), 464-477.
- [13] Axelrod, Robert, *The Evolution of Cooperation*, (New York: Basic Books, 1984).
- [14] Banfield, Edward C., *The Moral Basis of a Backward Society*, (New York : The Free Press, 1958).
- [15] Barkow, Jerome H., Leda Cosmides, and John Tooby, *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, (New York: Oxford University Press, 1992).
- [16] Baron, David P. , "Morally-Motivated Self-Regulation", Mimeo, (2007), Stanford University.
- [17] Barro, Robert J., and Rachel M. McCleary, "Religion and Economy", *Journal of Economic Perspectives*, 20 (2006), 49-72.
- [18] Becker, Gary S., "Nobel Lecture: The Economic Way of Looking at Behavior", *The Journal of Political Economy*, 101(1993), 385-409.
- [19] Becker, Gary S., *Accounting for Tastes*, (Cambridge, MA: Harvard University Press, 1996).
- [20] Ben-ner, Avner, and Louis Putterman, *Economics, Values and Organizations*, (Cambridge, U.K.: Cambridge University Press, 1998).

- [21] Bénabou, Roland, “Groupthink and Ideology”, Schumpeter Lecture at the meetings of the European Economic Association, *Journal of the European Economic Association*, (2008), forthcoming.
- [22] Bénabou, Roland, and Jean Tirole, “Belief in a Just World and Redistributive Politics”, *Quarterly Journal of Economics*, 121 (2006a), 699-746.
- [23] Bénabou, Roland, and Jean Tirole, “Identity, Dignity and Taboos: Belief as Assets”, Mimeo, (2006b), Princeton University.
- [24] Bisin, Alberto, Giorgio Topa, and Thierry Verdier, “Cooperation as a Transmitted Cultural Trait”, *Rationality and Society*, 16 (2004) , 477-507.
- [25] Bisin, Alberto, and Thierry Verdier, “A Model of Cultural Transmission, Voting and Political Ideology”, *European Journal of Political Economy*, 16 (2000), 5-29.
- [26] Bisin, Alberto, and Thierry Verdier, “The Economics of Cultural Transmission and the Dynamics of Preferences”, *Journal of Economic Theory*, 97 (2001), 289-319.
- [27] Bisin, Alberto, and Thierry Verdier, “Work Ethic and Redistribution: A Cultural Transmission Model of the Welfare State”, Mimeo, (2004), New York University.
- [28] Bisin, Alberto, and Thierry Verdier, “Cultural Transmission”, in *The New Palgrave Dictionary of Economics*. Second Edition, Steven N. Durlauf and Lawrence E. Blume, eds. (Palgrave Macmillan, 2008).
- [29] Bohnet, Iris, and Bruno S. Frey, “Social Distance and Other-Regarding Behavior in Dictator Games: Comment”, *The American Economic Review*, 89 (1999), 335-339.
- [30] Bowles, Samuel, and Herbert Gintis, “Social Capital and Community Governance”, *The Economic Journal*, 112 (2002), 419-436.
- [31] Cooter, Robert, “Expressive Law and Economics”, *Journal of Legal Studies*, 27 (1998), 585-608.

- [32] Cosmides, Leda, and John Tooby, "Cognitive Adaptations for Social Exchange" in *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, Barkow, Jerome H., Leda Cosmides, and John Tooby eds. (New York: Oxford University Press, 1992).
- [33] Davis, John, *Law and Family in Pisticci*, (London: London University Press, 1975).
- [34] Dawes, Robyn M., and Richard H. Thaler, "Anomolies Cooperation", *Journal of Economic Perspectives*, 2 (1988), 187-197.
- [35] Dhillon, Amrita and Jamele Rigolini, "Development and the Interaction of Enforcement Institutions", Mimeo, (2007), The University of Warwick.
- [36] Dixit, Avinash Kamalakar, "Trade Expansion and Contract Enforcement", *Journal of Political Economy*, 111 (2003), 1293-1317.
- [37] Dixit, Avinash Kamalakar, *Lawlessness and Economics - Alternative Models of Governance*, (Princeton: Princeton University Press, 2004).
- [38] Doepke, Matthias, and Fabrizio Zilibotti, "Patience Capital and the Demise of the Aristocracy", Mimeo (2005), University of California, Los Angeles.
- [39] Fehr, Ernst, Urs Fischbacher, and Simon Gächter, "Strong Reciprocity, Human Cooperation and the Enforcement of Social Norms", *Human Nature*, 13 (2002), 1-25.
- [40] Ferejohn, John, "Incumbent Performance and Electoral Control", *Public Choice*, 50 (1986), 5-26.
- [41] Fernandez, Raquel, "Alfred Marshall Lecture Women, Work, and Culture", *Journal of the European Economic Association*, 5 (2007a), 305–332.
- [42] Fernandez, Raquel, "Culture as Learning: the Evolution of Female Labor Force over a Century", NBER Working Paper no.13373, (2007b).
- [43] Fernandez, Raquel, "Culture and Economics" in *The New Palgrave Dictionary of Economics*. Second Edition, Steven N. Durlauf and Lawrence E. Blume, eds. (Palgrave Macmillan, 2008).

- [44] Fershtman, Chaim, and Uri Gneezy, "Discrimination in a Segmented Society: an Experimental Approach", *Quarterly Journal of Economics*, 116 (2001), 351-377.
- [45] Francois, Patrick, "Norms and the Dynamics of Institution Formation", Mimeo, (2006), University of British Columbia.
- [46] Francois, Patrick, and Jan Zabojník, "Trust, Social Capital & Economic Development", *Journal of the European Economic Association*, 3 (2005), 51-94.
- [47] Gambetta, Diego (ed.), *Trust: Making and Breaking Co-operative Relations*, (Oxford: Basil Blackwell, 1988).
- [48] Gambetta, Diego, *The Sicilian Mafia: The Business of Private Protection*, (Cambridge: Harvard University Press, 1993).
- [49] Giuliano, Paola, "On the Determinants of Living Arrangements in Western Europe: Does Cultural Origin Matter?", *Journal of the European Economic Association*, 5 (2007), 927–952.
- [50] Glaeser, Edward L. et al, "Measuring Trust", *The Quarterly Journal of Economics*, 115 (2000), 811-846.
- [51] Glaeser, Edward L., Rafael La Porta , Florencio Lopez-de-Silanes, and Andrei Shleifer, "Do Institutions Cause Growth", *Journal of Economic Growth*, 9 (2004), 271–304.
- [52] Greif, Avner, "Cultural Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies", *The Journal of Political Economy*, 102 (1994), 912-950.
- [53] Greif, Avner, "On the Inter-Relations and Economic Implications of Economic, Social, Political, and Normative Factors: Reflections From Two Late Medieval Societies" in *The Frontiers of the New Institutional Economics*, John N. Drobak and John V.C. Nye eds. (San Diego: Academic Press, 1997).
- [54] Greif, Avner, "Commitment, Coercion, and Markets: The Nature and Dynamics of Institutions Supporting Exchange", in *Handbook of New Institutional Economics*, Claude Ménard and Mary M. Shirley eds. (Berlin: Springer, 2005).

- [55] Guiso, Luigi, Paola Sapienza and Luigi Zingales, "Does Culture Affect Economic Outcomes?," *Journal of Economic Perspectives*, 20 (2006), 23-49.
- [56] Guiso, Luigi, Paola Sapienza and Luigi Zingales, "Social Capital as Good Culture," Marshall Lecture at the Meetings of the European Economic Association, *Journal of the European Economic Association*, (2008), forthcoming.
- [57] Hauk, Esther, and Maria Saez-Marti, "On the Cultural Transmission of Corruption", *Journal of Economic Theory*, 107 (2002), 311-335.
- [58] Herman, Benedict, Christian Thoni, and Simon Gächter, "Antisocial Punishment Across Societies", *Science*, 319 (2008), 1362-67
- [59] Ichino, Andrea, Fabian Bornhorst, Karl H. Schlag, and Eyal Winter, "Trust and Trustworthiness Among Europeans: South - North Comparison", EUI Working Paper, (2007).
- [60] Kaplow, Louis, and Steven Shavell, "Moral Rules, the Moral Sentiments, and Behavior: Toward a Theory of an Optimal Moral System", *Journal of Political Economy*, 115 (2007), 494-514.
- [61] Kranton, Rachel E., "Reciprocal Exchange: A Self-Sustaining System", *American Economic Review*, 86 (1996), 830-851.
- [62] Kumar, Krishna B., and John G. Matsusaka, "Village versus Market Social Capital", Mimeo, (2006), University of Southern California.
- [63] Levy, Gilat, and Ronny Razin, "A Theory of Religion: Linking Individual Beliefs, Rituals, and Social Cohesion", Mimeo, (2006), London School of Economics.
- [64] Lindbeck, Assar, and Sten Nyberg, "Raising Children to Work Hard: Altruism, Work Norms and Social Insurance", *Quarterly Journal of Economics*, 121, (2006), 1473-1503.
- [65] McMillan, John, and Christopher M. Woodruff, "Private Order under Dysfunctional Public Order", *Michigan Law Review*, 98 (2000), 2421-2458.



- [66] Massey, Douglas S., "A Brief History of Human Society: The Origin and Role of Emotion in Social Life", *American Sociological Review*, 67 (2002), 1-29.
- [67] Miguel, Edward, Paul Gertler, and David I. Levine, "Did Industrialization Destroy Social Capital in Indonesia ?", NBER working Paper (2002).
- [68] Miguel, Edward, and Daniel N. Posner, "Sources of Ethnic Identification in Africa", Mimeo, (2006), University of California, Berkeley.
- [69] Nisbet, Richard E., and Dov Cohen, *Culture of Honor - The Psychology of Violence in the South*, (Boulder: Westview Press, 1996).
- [70] Pinker, Steven, *How the Mind Works*, (New York: Norton, 1997).
- [71] Platteau, Jean-Philippe, *Institutions, Social Norms, and Economic Development*, (Harwood Academic Publishers & Routledge, 2000).
- [72] Polanyi, Karl, *The Great Transformation: the Political and Economic Origins of Our Time*, (Boston: Beacon Press, 1957).
- [73] Putnam, Robert, *Making Democracy Work: Civic Traditions in Modern Italy*, (Princeton University Press, 1993).
- [74] Tabellini, Guido, "Culture and Institutions: Economic Development in the Regions of Europe", Mimeo, (2005), Bocconi University.
- [75] Tabellini, Guido, "Institutions and Culture", *Journal of the European Economic Association*, 6 (2008), forthcoming.

# Figure 1

