Large populations can gain from economies of scale but lose internal trust due to diluted information. This creates an optimal group size. However, trusting strangers who claim to be members invites outsiders to disguise as insiders and abuse extended trust. Thus, if cultural diversity can raise the imitation cost it can promote cooperation. Even so, however, scale economies are lost when the population subdivides and the cultural boundaries may have to be enforced to prevent assimilation. The model is consistent with norms against inter-cultural marriage and episodic boundary-reinforcing conflict where formal institutions for contract enforcement are weak.

Keywords: Culture; Social Capital; Repeated Prisoners’ Dilemma.
JEL Classification: D82, D86, O17

1. Introduction
Abundant evidence suggests that trust is greater within than between culturally, racially, ethnically, or religiously homogenous groups (Knack and Keefer 1995, Easterly and Levine 1997, Alesina et al. 1999, Alesina and La Ferrara 2004, Horowitz 2001). For short, I will refer simply to cultural groups. While many theories explain in-group cohesion and out-group mistrust (see Horowitz 1998, 2001 for overviews), they usually assume a correlation between culture and some variable directly relevant for cooperation, such as interpersonal affinities or thick information flows among homogenous agents. While this may be true, I argue here that special relationships within the group are not required for inward trust and outward mistrust. More precisely, even if culture is merely a label containing no individual-specific information, and this fact is common knowledge, cultural diversity can promote trust. This finding contrasts with the common notion that diversity erodes cooperation and suggests that homogenizing factors such as nation building or globalization may be resisted.
I argue the following. First, trust between strangers without intrinsic mutual regard is only possible if current opportunistic behavior has future repercussions, which in turn requires information transmission (Kandori 1992). Second, the quality of information one can collect about the past of a partner is less in larger populations. Third, since populations sustaining trust will attract opportunistic outsiders pretending to belong, there must be a way to exclude non-members. Culture, however uninformative about a specific person, can be used for that purpose. To illustrate, suppose that a trader in Kenya meets a resident of Kenya and can try to check into his past before doing business. Even if she hears nothing bad, however, she may decide not to trust him because she only has time to ‘ask around’ in a small fraction of the country’s population. On the other hand, suppose that the trader meets a Swahili (one of Kenya’s tribes) and that the Swahili only do business with other Swahili. Now the trader only has to ask around in the Swahili subpopulation and is more likely to meet somebody familiar with the partner; hearing nothing bad is now a more reliable signal than before. If the resident reasons similarly, trust and trade between Swahili, but not between Swahili and others, may result. However, for this equilibrium to work a Swahili must be identifiable and each party must believe that the other party is no disguised pretender. This is ensured if Swahili culture is distinct and costly to imitate.

2. Model
There is a continuum population. In every period over the infinite horizon agents form pairs by mutual consent and play the prisoners’ dilemma stage game. No individual knows the history of play of her partner when they meet. However, after a pair is formed, but before playing the stage game, each party tries to learn whether or not the partner has cheated before. This is possible since cheating leaves memories or traces. When a player cheats, the probability of a future partner learning about it is \( q(x) \), \( q' < 0 \), where \( x \) is the measure of the population and for technical reasons I assume \( q \) is small.\(^a\) The normal form for the prisoners’ dilemma is shown below, where \( x, g, l > 0 \). Players discount the future by factor \( 0 < \delta < 1 \). The payoffs involving cooperation by at least one player are rising in population size \( x \) due to

\(^a\) For simplicity, the function \( q(x) \) is independent of the time passed after an instance of cheating. As time passes, on the one hand the news could spread more, but on the other hand it could fade away from records and memories.
economies of scale, such as knowledge sharing or specialization. “Cheating” means playing D when the partner plays C.

<table>
<thead>
<tr>
<th>P1/P2</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>$x, x$</td>
<td>$-1, x + g$</td>
</tr>
<tr>
<td>D</td>
<td>$x + g, -1$</td>
<td>$0, 0$</td>
</tr>
</tbody>
</table>

### 2.1. Equilibria

With perfect learning, the model works as if players carry labels which perfectly convey their individual past histories, as in Kandori (1992, section 5). Therefore, as in that paper, cooperation is possible if and only if it is possible with two players and perfect information. The necessary and sufficient condition with grim trigger strategies (the strongest possible punishment) is therefore $g \leq \frac{\delta x}{1 - \delta}$. On the other hand, with no learning at all no cooperation is possible. Proposition 1 identifies a cooperative equilibrium under partial information transmission.

**Proposition 1.** The strategy for each player “In every period, cooperate if and only if no past cheating is detected” is a sequential equilibrium in the repeated prisoners’ dilemma game with changing partners and imperfect information flows whenever

$$g \leq \frac{\delta x}{1 - \delta} - \sum_{t=1}^{\infty} \delta^t \left( \delta' \frac{g + x}{1 + qt} - \frac{\delta x}{1 - \delta} \right). \quad (1)$$

**Proof.** Sequential equilibrium refines perfect Bayesian equilibrium and applies if one can find, for each player, a sequence of totally mixed strategies converging to the player’s equilibrium strategy and associated beliefs following Bayes’ rule converging to the equilibrium beliefs (Kreps and Wilson 1982). Thus, since the proposed strategies yield zero probability of defecting, the only belief possible is that a player will meet no defecting players in the rest of the game regardless of her past experience. Consequently, we just have to check that a player is willing to cooperate when future partners will defect against her if and only if they learn that she cheated

---

The other, contagious defection equilibrium in Kandori (1992, section 4) is not possible here due to the infinite number of players.
in the past. First, because the rise in the proportion of future partners identifying the cheater declines with the number of cheating instances, it cannot be optimal to cheat a finite positive number of times. Second, a cheater’s expected net loss at time $t \geq 0$ after the first deviation is $p_t x + (1 - p_t)(-g)$, where $p_t$ is the likelihood of detection a time $t$ and $p_{t+1} = p_t p_r + (1 - p_t)(p_r + (1 - p_r)q) = p_t + (1 - p_t)^2 q$ . This is because with probability $p_t$ the cheater is discovered at time $t$, in which case both she and her partner defect, the number of cheatings stays the same and the likelihood of detection stay at $p_t$. Alternatively, with probability $(1 - p_t)$ the cheater goes undetected at $t$ and cheats again, in which case the proportion informed at $t + 1$ rises by the number of uninformed times the learning probability per uninformed $(1 - p_r)q$. Now using a continuous time approximation, which is reasonable for $q$ small, gives

$$\dot{p} = (1 - p)^2 q \Rightarrow \int \frac{dp}{(1 - p)^2} = \int q dt + C \quad \text{and using} \quad p(0) = 0 \quad \text{gives} \quad p_t = \frac{qt}{1 + qt} \quad \text{Thus,}$$

the present discounted net loss to cheating is non-negative if

$$\sum_{t=0}^{\infty} \delta^t \left\{ \frac{qt}{1 + qt} x - \frac{1}{1 + qt} g \right\} \geq 0 \quad \text{or (1). Informed agents will carry out the punishment since (D, D) is a Nash equilibrium in the stage game and only unilateral defection is punished by their future partners. No renegotiated agreement is feasible because, were the punisher to agree to cooperate after all, the cheater would still want to cheat.}$$

Perhaps more intuitive than the proof, the left hand side of (1) is the gain to defection in the current period and the right hand side is the net future loss, which is the present value of future cooperation foregone under perfect information (the first term) minus the ‘loss recovery’ due to the fact that information is imperfect: the fact that future partners may fail to learn of past cheatings allows the cheater to cheat more than once. The loss recovered per period if she is not detected is $x + g$ . The right hand side of (1) is strictly increasing in $q$, a measure of information quality, decreasing in the net gain to cheating, $g$, increasing in the discount factor $\delta$ and increasing in the mutual cooperation payoff $x$ holding $q$ constant. However, since in reality $q = q(x)$ declines as $x$ increases, a trade-off appears between the scale economies of group size and the dilution of information as the group size increases. I consider optimal group size below.
If the grim trigger strategies in proposition 1 cannot enforce cooperation, no strategies can. If grim trigger play suffices, however, one may suspect that milder and more efficient (off the equilibrium path) punishments may suffice. In one candidate alternative, - “cheat the cheater punishment” - the cheater cooperates for a finite number of periods after her defection and when she meets a player during that time who learns of her cheating this other player defects. The punishment time must be long enough to wipe out the gain to cheating one time in a present-value sense. For certain durations of punishment, this could be more efficient than grim trigger play, since the cheater suffers no greater loss and every punisher gains $x + g$ in comparison. Nonetheless, such strategies may be unlikely, because from the perspective of all his future partners a punisher would be indistinguishable from a cheater. Consequently, if agents have disincentives to cheat, they also have disincentives to punish, in which case they have incentives to cheat after all.\textsuperscript{c} On the other hand, however, rather than mutual defection in all future, the punishment could take the form of mutual defection until a finite time $T$ after every cheating, after which the cheater is forgiven and the players revert to cooperating. This requires that players can date detected instances of cheating and $T$ must satisfy, from the proof above, $\sum_{t=0}^{\infty} \delta^t \left( \frac{qt}{1+qt} x - \frac{1}{1+qt} g \right) \geq 0$.

2.2. Culture

Assuming existence, the payoff-maximizing group size is $x^* = \arg \max_x \frac{x}{1-\delta}$ subject to (1). However, a complication arises for any population measure $X > x^*$ since continuous defection would be a dominant strategy. Nonetheless, reality shows that often subgroups of a large population cooperate internally. To explain this, suppose now that every agent can be assigned membership of a single group and that members can identify each other as members. Assume also that any non-member of a group is treated by any member as if he were a member who cheated in the past. Then, if

\textsuperscript{c} If agents can learn both whether a partner ever cheated and whether that cheating was against a past defector, then “cheat the cheater” punishment may still be feasible. Alternatively, mixed-strategy equilibria with punishers indifferent between defecting and cooperating against a cheater may also exist. However, such an equilibrium would have mixed-strategy play spreading exponentially after a single defection and would not be robust to arbitrarily small uncertainty about utility functions (Bhaskar 1998).

\textsuperscript{1} Many ethnic groups themselves result from modernization and cultural assimilation (Horowitz 2001), but this does not diminish the threat from further assimilation.
membership is sufficiently costly to acquire (or imitate), group size can be restricted and cooperation enforced. In an equilibrium where membership can be acquired for free, but the number of members per group is restricted to $x^*$, out of $N+1$ groups the first $N$ would have membership $x^*$ and the last $x^* - Nx^* = \tilde{x} \leq x^*$.

Expecting defection, a group non-member has no incentive to approach a member.

**Example 1.** A population consists of 10000 racially white and 10000 racially black and otherwise identical agents. Let the prisoners’ dilemma payoffs be measured in thousands, $q(x) = 8/x^2$, and $\delta = 0.85$. Rewriting (1), cooperation requires $x/g \geq \sum_{t=0}^{\infty} \frac{1}{1+qt} / \sum_{t=0}^{\infty} \frac{qt}{1+qt}$. Computation shows that all 20000 agents can cooperate if $g \leq 2.04$, but intra-group cooperation, where blacks and whites only interact with others of the same race, can obtain for $g \leq 3.38$. Thus, with $2.04 < g \leq 3.38$, although the full scale economies of group size cannot be exploited, for payoff 20 per period, payoff 10 per period is possible if the racial diversity is exploited.

### 2.3. Group formation and cultural conflict

Suppose that permanent membership cards for different groups are issued once and for all when contract enforcement becomes an issue – due, perhaps, to population growth, quality decline of the information technology, changing payoffs in the stage game (e.g. economic growth), or the decline of third party enforcement (e.g. weak or failing states in many developing countries). Then group formation with internal trust can support intra-group cooperation indefinitely. In practice, such membership cards may be based on race, ethnicity, language, dialect, culture, or religion. This has the advantage that the relevant characteristic is (i) already possessed by some, but not all, agents in the population, (ii) to a large extent automatically transmitted across generations, through socialization, and (iii) costly to acquire for agents who do not have it initially.

However, in practice in the long run such group boundaries may become fluid as individuals are exposed to each other’s culture and intermarry, work, or attend school together. Also, there are direct incentives to assimilate since a person who can blend into many groups can obtain the unjustified trust of others and cheat them.
Outward assimilation by own group members and inward assimilation into one’s group are both problems. The consequent gains to enforcing group boundaries are consistent with occasional but severe group hostilities (Fearon and Laitin 1996, Leeson 2006), norms to prohibit inter-group marriage, and movements to preserve cultural autonomy and diversity in the face of nation-building and globalization. The model also fits the fact that such resistance efforts are typically provided at the group level, since assimilation may be individually rational. Proposition 2 gives sufficient conditions for group formation when members can assimilate at a cost and group sizes and assimilation costs may be asymmetric.

**Proposition 2.** The strategy for each player “In every period, cooperate if and only if no past cheating is detected and facing a same-group member” is a sequential equilibrium in the repeated prisoners’ dilemma game with changing partners and imperfect information flows for any population measure $X$ when a set of membership cards $C = \{c_1, c_2, \ldots, c_N\}$ exhausting the population exists such that the card acquisition cost function of every agent $z$ given by $K_z(c_m), m = 1, \ldots, N$, satisfies

(i) $K_z(c_n) \leq \frac{x(c_n)}{1-\delta}$ where $x(c_n) \leq x^*$ is the population measure joining group $n$ along with the agent; (ii)

$$
\sum_{m=0}^{N} I_m K_z(c_m) - K_z(c_n) \geq g - \frac{\delta}{1-\delta} x_{\min} + \sum_{r=t}^{x} \left\{ \frac{M(x_{\min} + g)}{M + q(x_{\max})} \right\}
$$

(2)

for all $\{I_1, I_{a-1}, I_{a+1}, \ldots, I_N\} \neq \{0, \ldots, 0\}$, where $I_m \in \{0,1\}$ is an indicator variable for membership in group $m$, $M = \sum_{m=0}^{N} I_m$, and $x_{\min}$ and $x_{\max}$ are the smallest and largest group sizes, respectively.

**Proof.** First, (i) says that every agent is willing to join her group. Second, (ii) says that the agent has no incentive to join any number of other groups and begin cheating. The cost of doing so is the left hand side of (2) and the net benefit is the right hand side. This is the left hand side of (1) minus the right hand side of (1) modified to reflect cheating in more than one group. If we first assume $x_{\min} = x_{\max}$, then (see the
proof of proposition 1) membership in M other groups gives the information diffusion equation \( \dot{p} = \frac{1}{M} (1 - p)^2 q \), because with symmetric group sizes a cheater should choose her partner in a given period from the group where her reputation is the least tainted (lowest likelihood of detection). Therefore, the rate of information spread in each group is slowed by a factor \( \frac{1}{M} \) and the derivation then follows the proof of proposition 1. Finally, using \( x_{\min} \) and \( q(x_{\max}) \) ensures that even agents coming from small groups (small gains to cooperation) who can join large groups (large gains to cheating) are not tempted.

**Example 2.** Suppose that, in a population of 30,000 agents, 10,000 have family cultural background labeled 1, 10,000 background 2, and 10,000 background 3. The cost of learning a culture for agent \( z \) is \( K_z(c_m) = b |I_m - I_z| \), where \( I_m \) and \( I_z \), \( m, z = 1, 2, 3 \), are indicators for the group the agent joins and the one her family comes from. \( b > 0 \) and a smaller \( b \) can be interpreted as more fluid cultural boundaries. Suppose functional forms and parameters as in example 1 and \( g = 3.38 \), so that \( x^* = 10 \) and efficiency obtains if every agent acquires her family’s culture at zero cost and cooperates every period within her group While condition (i) in proposition 2 clearly holds, however, we must also check condition (ii). Group 2 members are most likely to join other groups, so suppose a group 2 member buys membership of group 1 and begins cheating. This is unprofitable, however, for (rewriting (2))

\[
b \geq \sum_{t=0}^{\infty} \delta^t \left\{ \frac{0.08t}{2 + 0.08t} - \frac{2}{2 + 0.08t} \right\} \rightarrow 8.36.\]

Alternatively, she could join both groups 1 and 3 and begin cheating. However, this too is unprofitable as long as

\[
2b \geq \sum_{t=0}^{\infty} \delta^t \left\{ \frac{0.08t}{3 + 0.08t} - \frac{3}{3 + 0.08t} \right\} \rightarrow 12.12 \text{ or } b \geq 6.06.
\]

While in this example constrained efficiency obtains for \( b \) large, group formation generally may not be even constrained optimal and, due to membership acquisition costs, maximal size clubs (of size \( x^* \)) may not be optimal. For example, rather than change race and join a large group, it may be better not to change race and be in a small group.
Definition 1. Let the set of all possible card schemes satisfying (i) and (ii) in Proposition 2 be denoted $\Omega = \{C^1, C^2, ..., C^{n'}\}$. Scheme $C^h = \{c_1^h, c_2^h, ..., c_{N^h}^h\}$ divides the population into $N^h$ subgroups with $\int_{z\in n^h} z dz = x(c_n^h)$ members in group $n^h$. The socially optimal division scheme is

$$C^* = \arg \max_{C^h \in \Omega} \left\{ \sum_{n=1}^{N^h} \int_{z\in n^h} \frac{x(c_n^h)}{1-\delta} - K_z(c_n^h) dz \right\}. \quad (3)$$

For an example of a suboptimal outcome, suppose that a population of measure one can divide either into religious groups of measure $1/2, 1/4$ and $1/4$ or into ethnic groups of measure $1/3$ each. Assume that the latter is efficient but for some reason, perhaps due to low joining costs, the first ethnic group and half of the second ethnic group end up forming the large religious group. If the third ethnic group were still to form, a member could join the religious group and cheat in $4/3 > 5/6 > 3/4$ of the population, which could be tempting. A constrained inefficient religious-divisions equilibrium may then arise.

3. Conclusion

This paper has explained cultural diversity as a means to sustain cooperation when third party contract enforcement and information flows are weak and information quality declines in the size of a cooperating group. In equilibrium, agents sharing a culture cooperate precisely because for this reason. Provided that it is distinct and costly to imitate, the only information culture needs to convey is membership. Depending on the payoff structure of the game, the information technology, and the distribution of observable ex-ante heterogeneity in the population, different size distributions may appear and may not be constrained optimal. Improved cheating detection (information) technology can raise efficiency via economies of scale and limit the need for cultural diversity and perhaps for periodic boundary reinforcing group conflict. This may explain why ethnic divisions, for example, appear particularly relevant to understand behavior in developing countries.
Acknowledgements

I am grateful to participants at The Association for the Study of Nationalities Annual Conference, Columbia University, New York, NY, April 12-14 2007 and, in particular, Klaas T. van’t Veld for comments.

References