Facoltà di Economia di Forlì
Laurea Magistrale in Economia Sociale

Working Papers
102

Cultural Diversity, Cooperation, and Anti-social Punishment

Marco Faillo
University of Trento
Daniela Grieco
Bocconi University and University of Verona
Luca Zarri
University of Verona

Giugno 2012

Info: **AICCON** - Tel. 0543.62327 - ecofo.aiccon@unibo.it - www.aiccon.it

Cultural Diversity, Cooperation, and Anti-social Punishment

Marco Faillo[1] Daniela Grieco[2] Luca Zarri[3]

Mar 12, 2012

Abstract

Is culture an important variable to explain whether groups can successfully provide public goods? A wealth of empirical evidence on both industrialized and developing countries shows that cooperation levels decrease in the presence of ethnic divisions. Although several laboratory works deal with cultural differences, so far most studies restrict their attention to cross-cultural comparisons among internally homogeneous societies. We depart from these contributions and conduct an intercultural public goods game with punishment experiment in Italy, a country where immigration is a quite recent, but politically hot phenomenon. We investigate the effects of introducing a varying number of foreign participants within a homogeneous pool of native subjects. Our results indicate that foreigners contribute significantly less than natives, natives react lowering their own contribution levels, and, consequently, the degree of cultural diversity negatively affects the overall level of cooperation. In terms of sanctioning, we observe no difference in the overall amount of assigned and received punishment points; however, behaving mostly as free-riders, foreigners are more likely to use anti-social punishment. In the absence of institutional restrictions ruling out anti-social punishment, this might amplify the documented detrimental effect on cooperation.

JEL Classification: C72; C91; C92; D64 ; D71.

Keywords: Experimental Economics; Public Good Games; Cooperation; Cultural Diversity; Anti-social Punishment.

1. Introduction

[1] marco.faillo@unitn.it, Department of Economics, University of Trento.
[2] daniela.grieco@unibocconi, Department of Economics, Bocconi University and University of Verona.
[3] luca.zarri@univr.it, Department of Economics, University of Verona.

All over the world, social and economic interaction is shaped by cultural factors. In particular, it often takes on an ethnic character. Groups frequently mobilize along ethnic lines for civil war (Cederman and Girardin, 2007) and ethnic divisions may impact the formation of political parties and the distribution of public resources (Chandra, 2004). Further, economic activities might take place in ethnic enclaves (Fafchamps, 2000), especially under asymmetric information (Greif, 1993). Ethnic fragmentation turns out to be correlated with national government corruption and low economic growth (Mauro, 1995; Easterly and Levine, 1997), lower participation in voluntary or civic organizations (Alesina and La Ferrara, 2000) and lower levels of trust (Alesina and La Ferrara, 2002). A large and growing empirical literature relates high levels of ethnic diversity to low levels of public goods provision (e.g., Alesina and Levine, 1997; Alesina et al., 1999; Khwaja, 2002; Miguel and Gugerty, 2005). Such negative correlation emerges both within industrialized and developing countries, to the point that Banerjee et al. (2005) describe it as "one of the most powerful hypotheses in political economy" (p. 639). Despite the availability of a wealth of empirical evidence, however, the specific channel(s) through which this relationship operates remains poorly understood. *Why*, then, does ethnic diversity undermine the provision of public goods? Why is public goods provision more successful in homogeneous communities?

The advantage of economic experiments is the higher degree of control allowed by the lab, compared to the field. So far, however, most laboratory studies have restricted their attention to culturally homogeneous groups formed by subjects sharing the same nationality (see e.g. Roth et al., 1991; Henrich et al., 2001): the key methodological feature of such *cross-cultural* studies is that they are typically based on a comparative approach, with the same experiment being run in a number of (culturally) different locations and a subsequent comparison of the results obtained in each location. By contrast, only a few experimental studies have been explicitly targeted at *intercultural* environments, i.e. heterogeneous groups formed by subjects from different societies. As noted by Ferraro and Cummings (2007), observations of these laboratory societies might offer insights about behavior and outcomes in the naturally-occurring societies outside the laboratory. In particular, we claim that a significant advantage of creating an intercultural environment in the lab is that it makes it possible to look at how subjects from *different* countries interact within a common social environment and play the *same* game. Our work falls in this second category, as we conducted a public goods game experiment in which sessions differed in terms of participants' country of origin. Our experiment was run in Italy, a country where immigration is a quite recent, but politically hot phenomenon and only first-generation immigration is common at the moment. We manipulated the degree of cultural and ethnic heterogeneity in the experimental sessions, that differ in terms of participants' country of origin: our subjects were mainly Italian, but a sizeable minority of participants was from African and Eastern Europe countries. Hence, we focused on an intercultural but relatively homogeneous environment, where the majority of subjects can be viewed

as ‘natives’ and the minority as ‘foreigners’. An important point to be made is that the growing literature on culture not only contributes to explain differences in economic performance across different culturally homogeneous societies (as Bornhorst et al. (2010) observe), it also helps understand differences in economic performance between heterogeneous societies characterized by different degrees of cultural diversity *within* themselves.

Our findings indicate that natives and foreigners do differ in cooperation levels from the outset and that the degree of cultural diversity significantly lowers both natives’ and foreigners’ contribution levels in the game. Next, our analysis illuminates a mechanism relevant to understanding this negative correlation, that is the role of *peer punishment* in the game. The structure of the remainder of the paper is as follows. Section 2 reviews the related empirical literature on cultural and ethnic diversity and cooperation, including references to laboratory studies. Section 3 contains an illustration of our experimental design and procedure. Section 4 outlines the major results and Section 5 concludes.

2. Related literature

A rich empirical evidence from the field shows that public goods like years of schooling, school funding and facilities, health care, social security, paved roads, the efficiency of the electricity network, sewers, trash pickup, water provision and water well-maintenance are inversely related to ethnic fragmentation (Alesina and Levine, 1997; Alesina et al., 1999; Miguel and Gugerty, 2005). Charles and Kline (2006) find that greater ethnic homogeneity across American neighborhoods is associated with higher rates of car-pooling to get to work. Vigdor (2004) documents a similar relationship by considering mail responses to the 2000 Census in the United States. As he observes, to date researchers have devoted little effort to providing an economic motivation for ethnic fragmentation effects. Despite the robustness of this result with regard to both advanced and developing economies, a crucial question remains largely open: *why* do we observe this negative correlation between ethnic heterogeneity and public goods provision?

Existing work on the issue suggests that a number of distinct mechanisms might be simultaneously at work. First, different ethnic groups may have different *preferences* about the type and size of the public good to be financed (Alesina et al., 1999; Alesina and La Ferrara, 2005). Second, a group’s utility deriving from getting access to the public good might be reduced when a different ethnic group also uses it. As Vigdor (2002) notes, an individual’s willingness to pay for a local public good or service derives partly from the fact that the benefits that the public good confers on other members of her community enter directly into her utility function. However, the individual might differentially value those benefits depending on their being conferred to members

of her own ethnic group or to members of different ethnic groups[4]. In other words, subjects may display parochial preferences. Third, heterogeneous groups show some inability to impose social *sanctions* (Miguel and Gugerty, 2005). A fourth channel which has been identified in the empirical literature is *trust*. Banerjee et al. (2005) emphasize that in market economies even in non-war times "trade often requires trust, providing public goods needs collective action and the rule of law is only possible if everyone accepts the rights of others" (p. 639). Bahri et al. (2005) refer to trust in order to account for the presence of higher cooperation in homogenous societies: higher trust has been associated with greater citizen involvement in politics, lower corruption, more effective public services, higher economic growth. Similarly, Alesina and La Ferrara (2000, 2002) show that heterogeneous local communities exhibit lower levels of trust and of other components of social capital: ethnic differences determine the rise of more particularized trust, where confidence is restricted to a narrow set of ingroups like family and friends.

In terms of empirical methods, it is worth noting that while a large number of field works document a negative association between ethnic heterogeneity and cooperation, only a few experimental papers specifically deal with this relationship. As we anticipated in the previous section, a plethora of studies confine their attention to cross-cultural comparisons. Roth et al.'s (1991) pioneering analysis is based on the same two-person bargaining and multiperson market setups being conducted with four culturally and geographically distant subject pools in the United States, Yugoslavia, Israel and Japan. Burlando and Hey (1997) compared British vs. Italian subjects' behavior within a repeated public goods game framework. Ockenfels and Weimann (1999) focused on East vs. West Germans and Brosig et al. (2011) repeated their solidarity experiment 20 years after the fall of the Berlin wall. Henrich at al. (2001) wondered whether the same qualitative results obtained with Western subjects hold also when participants come from small-scale societies in developing countries. Cason et al. (2002) concentrated on American vs. Japanese players[5].

Recently, Gächter and Herrmann (2009) compared Russians' vs. Swiss' behavior, including their willingness to punish others at a cost. Beyond the social dilemma framework, Fershtman and Gneezy (2001) were interested in looking at Israeli subjects from different ethnic groups and Eckel and Petrie (2011) compared white and nonwhite subjects playing a trust game. A common feature in all these experiments is that each session only includes participants from a single group, i.e. the comparison was made among homogeneous environments. In contrast, Ferraro and Cummings (2007) conducted an ultimatum game experiment designed to determine whether cultural diversity affects subjects' decisions. They varied the ethnic mix of bargaining sessions and studied

[4] He adds that Luttmer's (2001) model is based on a similar intuition: here, agents derive utility (or disutility) from the outcomes of others, but different people are attached different weights in the utility function.
[5] See also Brandts et al.'s (2004) four country comparison.

intercultural interactions among members of Hispanic and Navajo cultures in the United States. Their findings indicated that Hispanic and Navajo subjects behaved differently and that their behavior is influenced by the ethnic composition of the experimental session. An interesting feature of their design is that they explored behavioral variability in response to changes in the proportions of the two cultural groups in an experimental session. Specifically, they focused on three contexts: (1) all players share the subjects' culture; (2) the player's culture makes up a large majority of the players and (3) the players' culture is a small minority of the players. On a similar vein, Finocchiaro Castro's (2008) lab experiment investigated the effects of interactions between Italian and British subjects by comparing the behavior of groups formed by subjects of the same nationality but also considering the behavior of mixed groups composed of half British and half Italian subjects. His results showed no country effect, but significantly lower individual contribution levels in heterogeneous groups. Carpenter and Cardenas (2011) allowed participants from different cultures (Colombia and the United States) to play a common pool resources game by interacting in real time within the same session and to make decisions within the same experimental context.

Bornhorst et al. (2010) run an experiment in which students of different European nationalities are matched in groups of five and repeatedly choose with whom within their group they want to play a trust game. Participants observe age, gender, nationality and number of siblings of the other players. The region of origin, "North" or "South" is a major determinant of success in the experiment. Moreover, participants tend to trust those they trusted before and who trusted them. They do not find evidence of regional discrimination per se. It is only the underlying and significant differences in behavior that translate through repeated interactions into differences in payoffs between the two regions.

By means of a third-party punishment experiment, Hoff et al. (2011) investigate how the social structure of a society affects individuals' willingness to sanction norm violations, by focusing on how the assignment to the top or bottom of an extreme hierarchy – the Hindu caste system– affects individuals' willingness to punish violations of a cooperation norm. They find that low caste punish norm violations less often and less severely than high caste individuals, even controlling for wealth, education and participation in village government.

Finally, it is worth mentioning Habyarimana at al.'s (2007) work, where a set of experimental games, typically played in laboratory environments, has been run in Kampala, Uganda, a place where ethnicity is highly salient in everyday social interactions. In their experiment, the information on other players' ethnicity is critical, and treatments differ as subjects might interact face-to-face or using a computer interface where images of other players could be made available. The aim of that work was to disentangle plausible explanations about why ethnic homogeneity favors public goods provision: commonality of tastes, common cultural material (language, experience, modes of interaction) that makes it easier for community members to

communicate and work together, and strategy selection mechanisms in settings with multiple equilibria where ethnic groups possess both norms and networks that facilitate the sanctioning of community members who defect.

3. An intercultural public goods game experiment

3.1. Experimental design

The experiment consists of a finitely repeated voluntary contribution mechanism (VCM) with punishment options that is grounded on Fehr and Gächter's (2000) seminal work. In each of the 20 periods, participants take part into a two-stage game: at stage 1, players simultaneously choose how much to contribute to the public good (contribution stage) and at stage 2 they have access to punishment options (punishment stage). We have recourse to two treatments (Baseline and Restricted) that differ for the presence of 'virtuous' restrictions to sanctioning opportunities: whereas in the Baseline each participant can punish any other peer in the group, regardless of her contribution behavior at stage 1, in the Restricted treatment a subject is entitled to punish another subject at stage 2 only if her contribution at stage 1 has been higher than the contribution of the peer she wants to punish (see Faillo et al. (2010) for a paper with an explicit focus on this form of 'legitimate punishment'). The presence of virtuous restrictions rules out by construction the possibility to observe anti-social punishment (Herrmann et al., 2008). Like in a standard, finitely repeated VCM, insofar as all the subjects are supposed to be driven by material self-interest only and this information is common knowledge, the unique subgame perfect equilibrium is for all agents to free ride in every period and never punish.

Three key features of our experimental environment can be described as follows. As we anticipated in the previous section, our major goal was to shed light on the relationship between cultural diversity and cooperation by exploiting the higher degree of control made possible by having recourse to the experimental laboratory rather than to other empirical methods in the field.

First, in a public goods game framework, the existence of a single public good to be voluntarily provided allows us to rule out that underprovision depends on different groups having diverging tastes about which specific public goods should be provided (Alesina et al., 1999; Alesina and La Ferrara, 2005). Given our research question, we view this as an important advantage of the experimental methodology, as in the field it is often extremely difficult to rule out that a negative association between diversity and cooperation depends on taste divergence over the type and size of the public good (e.g. school or water maintenance) to be provided among different groups.

Second, unlike several existing studies, we investigate this relationship within a relatively homogeneous society. Our experiment has been conducted in Italy, which is the country where all the components of our subject pool live, but where Italians are unambiguously identifiable as the 'natives' and Non-Italians as 'foreigners'. As noted by Reuben and Riedl (2009), norms have been found to influence behavior even *after* people have moved across societies: for example, norms strongly impact U.N. diplomats' illegal parking behavior in New York (Fisman and Miguel, 2007) and the level of trust exhibited by descendents of immigrants to the U.S. is correlated with the degree of trust of their ancestors' country of origin. On a similar vein, Guiso et al.'s (2006) findings are consistent with the idea that priors have a cultural component, which is transported to the new world and continues to affect individual beliefs even in the new environment and even several generations later. In their influential paper on social capital and financial development in Italy, Guiso et al. (2004) showed that the behavior of movers is still affected by the level of social capital of the province where they were born (see on this also Brosig et al., 2011). Hence, cultural values seem to affect one's behavior even when the person moves and even after long periods of time. In our experiment, since our subject pool is made of university students, our foreigners are people who only recently arrived in Italy in order to enroll as undergraduate students. Therefore, it is very likely that these participants' values, beliefs and behavioral choices are affected to a significant extent by their own ethnic and cultural background. Since we hypothesize that a similar dependence on one's ethnic and cultural background holds for natives, it is possible to argue that our intercultural setup resembles to some extent a natural experiment.

Next, a third crucial feature of our design is that we purposely provided no cues about coplayers' nationality to the subjects: even though our major research question was about the behavioral relevance of participants being either natives or foreigners, we *never* made the existence of such 'groups' *salient* to the players. In other words, our subjects never knew, neither before, nor during nor after the experiment that we were interested in focusing on the potential impact of the cultural dimension on their choices throughout the game. In doing this, we significantly departed from some of previous experimental studies (where participants were provided with pictures, surnames or direct information about the nationality of their coplayers; see e.g. Fershtman and Gneezy (2001) and Bornhorst et al. (2010)) and avoided to make subjects' nationality salient in our treatments. In this regard, we claim that we went even beyond the so well-known 'minimal group paradigm', where membership was made salient through arbitrary and virtually meaningless 'labels' signaling group membership (see e.g. Frank and Gilovich, 1998). In our experiment, we were totally neutral on this, so that in our setup we view salience of culture as even more minimal than in the classical minimal group studies. By so doing, we can rule out that the results we obtained (see Section 4) can to some extent depend on experimenters' *demand effects* (see on this Zizzo, 2010).

3.2. Experimental procedure

A total of 152 subjects participated voluntarily in the experiment at the CEEL Lab of the University of Trento. A total of 6 sessions were conducted, between March 2010 and October 2011. Six sessions had 20 participants and the other two had 16 participants. Part of the data have been used and described in our companion paper on 'legitimate punishment' (Faillo et al., 2010). The experiment was programmed by using the z-tree platform (Fischbacher, 2007). The subjects were undergraduate students (60% from Economics, 48.7% females, 80% Italian). No individual participated in more than one session. In each session, the participants were paid a 5 euro show-up fee, plus their earnings from the experiment. The average payment per participant was 14.4 euros (including the show-up fee) and the sessions averaged approximately 1 hour and 30 minutes. Prior to each session, participants met in a common room, where they were welcomed by the experimenter. By doing this, we allowed subjects to observe the other participants taking part in the same session. After few minutes, they were led to the lab by the experimenter and asked to draw lots, so that they were randomly assigned to terminals. Once all of them were seated, the instructions were handed to them in written form before being read aloud by the experimenter. We took great care to ensure that the participants understood both the rules of the game and the monetary incentives. They had to answer several control questions and we did not proceed with the actual experiment until all participants had answered all questions correctly.

In each session, there were 20 periods of interaction that proceeded under identical rules. The participants in a session were randomly assigned to groups of size four, so that they did not know the identities of the other members of their group. Like other experimental studies (see e.g. Cinyabuguma et al., 2006; Denant-Boemont et al., 2007), we used a partner protocol that kept the composition of each group constant over rounds, so that, at the end of each period, individuals remained in the same group. However, individuals' labels were reassigned on a random basis in each period. For example, the same player could be designated as player 45 in period t, as player 6 in period t + 1, and as player 38 in period t + 2. Therefore, our partner protocol was also characterized by anonymity of the components of the group and change of participants' labels across rounds. The design and the parametric structure of the experiment are based on Fehr and Gächter (2000).

3.3. Treatments

We implemented two experimental treatments: a baseline, unrestricted punishment and full information (Baseline) treatment, and a restricted punishment with full information (Restricted) treatment. There were 4 sessions (20 subjects in three sessions and 16 in the other) for the Baseline, and 4 sessions (with 20 subjects in three sessions and 16 in the other) for the Restricted. For each treatment, in each session the subjects were divided in groups of N=4 (in line with many VCM experiments) subjects, who played a two-stage finitely repeated public goods game with punishment options for T=20 periods. Participants were aware of the number of rounds they were going to play and of the number of stages: information on the following stages permits to evaluate the effect of the threat of being punished at stage 2 and on contribution decisions at stage 1.

At stage 1, at the beginning of each period each participant receives a fixed amount e = 20 of tokens. Each participant i has to decide whether she wants to invest into a public project or not an amount $g_i \leq e$. Decisions are made simultaneously and with no information about peers' choices. At the end of stage 1, each participant is informed about her current earnings, which consist of two components:

a. The proportion of her initial 20 tokens that she has kept for herself (i.e. 20 tokens – Her Contribution to the project);
b. Her income from the project. The income to her is equal to 40% of the total of the four individual contributions to the project.

Therefore, her earnings at the end of stage 1 are calculated by the computer in the following way:

Each participant's earnings after stage = (20 – her contribution to the project) + 40%(total group contribution to the project)*

$$= (20 - g_i) + 0.4\sum_{j=1}^{4} g_j$$

In the Baseline treatment, any participant has access to stage 2, i.e. there are no limitations over who can punish and who can be punished. By contrast, in the Restricted treatment participants know that they can go on with stage 2 only if they contribute *more* than their peers. Specifically, player i will be entitled to sanction player j at stage 2 only if $g_i > g_j$. At stage 2, subjects are given the opportunity to simultaneously punish those who contributed less than them by assigning a certain amount of points. This implies that the highest contributor in a group is fully immune from punishment. Potential punishers might decide to assign up to 10 points to each punishee: point assignment is costly and costs are charged according to a standard convex cost function as in Fehr and Gächter (2002) (see Table 1).

[TABLE 1]

Each point that a subject receives reduces her earnings at stage 1 by 10%.

Each participant's earnings at the end of stage 2 are calculated by the computer in the following way:

Each participant's earnings after stage 2 = earnings at the end of stage 1- cost of points she assigned at stage 2 - 10% number of points received*earnings at the end of stage 1*

4. Results

Table 2 summarizes the number of Italian and foreign (specifying their own nationality) subjects in all the experimental sessions.

[TABLE 2]

Table 3 relates cultural diversity in each session to subjects' average contribution levels and summarizes natives' and foreigners' contribution decisions in the first period and in all the periods.

[TABLE 3]

In the first period, foreigners' average contributions can be lower or higher than natives', depending on the session, in both treatments. However, if we look at all the 20 periods of interaction, foreigners always exhibit a lower level of average contribution, with the exception of Session Restricted7. For natives, restrictions appear to play a crucial role in shaping contribution levels in the first period (Column 1) and in the whole 20 periods (Column 4): in the presence of restrictions, natives' contribution levels are always higher than without restrictions. This is not true (see e.g. sessions Baseline1 and Restricted5) for foreigners. The difference in the level of average

contribution between natives and foreigners in both treatments is summarized in the following Figure.

[FIGURE 1]

[FIGURE 2]

As far as foreigners' attitude and contribution/punishment behavior in our public goods game environment are concerned, we offer two competing hypotheses. On the one hand, it is plausible to conjecture that having natives and foreigners interacting and playing the same game in order to deal with a collective action problem may prompt reactionary choices on the part of the latter group (but possibly also on the part of the former group) that reinforce parochial distinctions among people. In other words, taking part in a social dilemma where the provision of a single, *common* public good is involved may reinforce *parochialism* by strengthening the demarcation between one's group and the outgroup. On the other hand, we may hypothesize that since our foreigners are young people who voluntarily went to live and study in a different country, the intercultural environment will weaken the relevance of nationality as a source of identification and broaden the boundaries of the group to which individuals perceive they belong[6].

Though in principle we view both hypotheses as plausible, we believe that the first one is more in line with empirical evidence from ethnographic and socio-economic studies. In light of this, the major goal of our intercultural experiment is testing whether foreign subjects do actually contribute less than non-foreign subjects, by shedding light on the role of punishment in driving participants' behavior in the game.

In the aim of exploring the influence of subjects' nationality, which, in line with previous work in the area (see e.g. Bornhorst et al., 2010), we view as a proxy for 'culture',, we estimate the following equation:

$$contribution_{it} = \beta_0 + \beta_1 nat_i + \beta_2\, restricted \quad + \beta_3 subperiod + controls \quad (Eq.1)$$

Results are summarized in Table 4.

[6] For an experimental paper testing two similarly stated hypotheses, though with a specific focus on the role of 'globalization' in affecting people's attitudes towards close vs. distant others and propensity to cooperate, see Buchan et al. (2009).

[TABLE 4]

.

Result 1. *Foreigners' contribution levels are significantly lower than natives' ones.*

Regardless of the presence of institutional restrictions, nationality is always significant in affecting contribution levels. Two possible interpretations of this first result can be offered: (1) a 'country effect' occurs in shaping contribution levels, i.e. Italians are more cooperative than foreigners; (2) there is a 'majority vs. minority group effect', if natives (the majority group) and foreigners (the minority group) react differently to the degree of cultural heterogeneity they experience in the game (see below). Foreigners come mainly from African and Eastern Europe countries and the difference in contribution levels between these two sub-groups of foreign participants is not significant: regressions have been replicated by adding the dummy 'African'.

In the presence of virtuous restrictions, the overall contribution level rises, but the difference between Italians and non-Italians remains. This can be explained by the fact that anti-social punishment has a negative effect on the level of contribution of the victim (see below, Table 5). We also conjecture that in the Restricted treatment high-contributors know from the beginning that they cannot be punished by participants who contribute less than them, and this might account for the higher average contribution observed already in early periods in the Restricted treatment.

Result 2. *Anti-social punishment has a detrimental effect on contribution levels.*

The regression described above is modified by adding as further regressor the number of anti-social punishment points received by the subject in the previous period. Results are summarized in Table 5.

[TABLE 5]

In the absence of restrictions, high-contributors (who are mainly natives, as shown in Result 1) can be victims of anti-social punishment. Let us now turn to investigate the determinants of anti-social punishment. In particular, we aim at analyzing the percentage of anti-social punishment points distributed by natives and foreigners, respectively.

Result 3. *Given the contribution level, natives are not punished anti-socially more than foreigners*.

As high-contributors, natives do not receive anti-social punishment more frequently, but this happens just because they contribute significantly more: there is no role for nationality *per se* in shaping the attitude to anti-socially punish others, and the level of contributions significantly and negatively affects the likelihood to receive anti-social punishment points. See Table 6 for more detailed results.

[TABLE 6]

Table 6 reports the percentage of received anti-social punishment points: as punishment options are limited to pro-social points in the Restricted treatment, the table refers to the Baseline only.

Summing up, experimental sessions with a higher number of foreign subjects show a significantly lower level of cooperation because (1) foreigners contribute significantly less, and (2) foreigners use more frequently anti-social punishment but only as lower contributors,. When virtuous restrictions against anti-social punishment are at work, the latter effect is ruled out, and cooperation is higher, but still lower compared to the levels reached in more homogeneous sessions.

Explanations of this finding can be related to two main mechanisms. Foreigners might contribute less due to cultural traits that can be more or less prone to individualism and competition: this has been shown by Gaechter and Hermann (2011) as a typical feature of former communist countries. Alternatively, restraining from free riding and contributing to a public good may turn out to be a 'luxury good', so that lower cooperation would be explained by lower income (see also Bornhorst et al. (2010) for a similar interpretation with regard to their trust game environment). Data on students in the university of Trento show that, on average, foreigners are poorer for any income range (see Figure 3). We aim at investigating the importance of income further by collecting income data on our subject pool, to see whether it plays a role, possibly jointly with subjects' nation-specific cultural factors, in driving our result.

[FIGURE 3]

5. Discussion and concluding remarks

In our public goods game experiment, cultural diversity has a detrimental effect on cooperation both when we consider the whole sample and under unrestricted punishment only (baseline treatment). Further, what reduces cooperation is the number of foreigners in the session. Next, natives contribute more than foreigners to the public good to be provided.

The degree of control provided by the experimental lab allowed us to shed light on a specific channel through which the above relationship operates, so that we can contribute to answer the question concerning the reasons *why* cultural diversity negatively impacts public goods provision. In particular, our data suggest that this correlation passes through the role of peer punishment.

In this regard, we showed that the relationship between culture and the (unrestricted vs. restricted) nature of punishment is important. In particular, in sessions where anti-social punishment is possible, sanctioning activity appears to be ineffective in sustaining cooperation, and this result is stronger the higher the number of foreigners. However, our results also indicate that while under an unrestricted punishment regime cultural diversity negatively affects public good provision, restricted punishment favors the enforcement of cooperation despite the presence of cultural diversity. The reason why this occurs is that restricted punishment (a) rules out a quantitatively significant phenomenon such as antisocial punishment[7] by construction and (b) exerts a differential impact on natives' and foreigners' behavior, by inducing natives and foreigner high-contributors to contribute more.

Hence, it is worth noting that on the whole not only our findings illuminate some specific mechanisms driving the emergence of a negative relationship between cultural diversity and cooperation, they also identify a possible solution for the problem under study. Our analysis revealed that a sanctioning institution based on legitimacy, that is an institutional arrangement where peer punishment, far from being unrestricted, depends on a principle of legitimacy (Faillo et al., 2010), is effective in preventing the negative impact of cultural diversity on cooperation. This result suggests that even culturally heterogeneous communities can be successful in providing public goods insofar as they rely on shared rules based on a common, socially acceptable sanctioning rule. When an institution is grounded on fair principles, it is more likely to expect that even people of different origin will agree on its implementation and act accordingly. Therefore, it seems that people can successfully overcome the 'ingroup-outgroup' tension of parochialism and

[7] Anderson and Putterman (2006): find inter alia that a surprisingly large share of punishment events involve low contributors punishing high ones (p. 4).

experience a sense of common belonging merely by virtue of the adoption of a fair and shared principle of legitimacy.

In concluding their paper on culture and economic outcomes, Guiso et al. (2006) pointed out that important cultural elements "will make economic discourse richer, better able to capture the nuances of the real world, and ultimately more useful" (p. 46). They also raised an exciting set of questions, among which there was the following: what is the interaction between culture and formal institutions? Our work sheds light on this issue, thanks to the high degree of control allowed by the experimental lab: we compared two punishment regimes and found that culturally different individuals react differently to them. In particular, we showed that when both natives and foreigners are present, a legitimate punishment institution works better than a standard, unrestricted punishment mechanism.

ACKNOWLEDGEMENTS: We gratefully acknowledge the Italian Ministry of Education, University and Research (2008 Prin project on "Social Capital, Corporate Social Responsibility and Performance" – Padova unit) and the Departments of Economics of the Universities of Trento and Verona for financial support. Our gratitude also goes to Cristina Bicchieri, Alessandro Bucciol, Pierluigi Conzo, Luca Corazzini, Giorgio Coricelli, Paola Dongili, Tore Ellingsen, Francesco Farina, Enrique Fatas, Antonio Filippin, Gianluca Grimalda, Diego Lubian, Laura Magazzini, Charles Noussair, Daniel Zizzo and to the participants in the Prin 2008 Bologna meeting (Apr 2011), the XIV Summer School on Economics and Philosophy in San Sebastian (Jul 2011), the European ESA Conference in Luxembourg (Sep 2011), the International Conference on "Socially Responsible Behavior, Social Capital and Firm Performance", University of Milano-Bicocca, Milan (Oct 2011), the SEET Meeting in Lanzarote (Feb 2012), the IMEBE meeting in Castellon (Mar 2012) and the colleagues of the Department of Economics of the University of Padua for useful suggestions and comments. The usual caveats apply.

References

Abbink, K., Brandts, J., Herrmann, B., Orzen, H. (2010). Intergroup conflict and intra-group punishment in an experimental contest game. American Economic Review, 100, 420-447.

Alesina, A., Baqir R., Easterly, W. (1999). Public goods and ethnic divisions. Quarterly Journal of Economics, 114, 1243-1284.

Alesina, A., Fuchs-Schündeln, N. (2007). Good-bye Lenin (or not?): The effect of Communism on people's preferences. American Economic Review, 97, 1507-1528.

Alesina, A., La Ferrara, E. (2005). Preferences for redistribution in the land of opportunities. Journal of Public Economics, 89, 5-6, 897-931.

Banerjee, A., Iyer, L., Somanathan, R. (2005). History, social divisions and public goods in rural India. Journal of the European Economic Association, 3, 639-647.

Bernhard, H., Fischbacher, U., Fehr, E. (2006). Parochial altruism in humans. Nature, 442, 912-915.

Bornhorst, F., Ichino, A., Kirchkamp, O., Schlag, K., Winter, E. (2010). Similarities and differences when building trust: the role of cultures, Experimental Economics, 13, 3, 260-283.

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

Boyd, R., Richerson, P.J. (1995). Why does culture increase human adaptability? Ethology and Sociobiology, 16, 125-143.

Boyd, R., Richerson, P. (2009). Culture and the evolution of human cooperation. Philosophical Trans R Soc Lon B Biol Sci, 364, 3281-3288.

Brandts, J., Saijo, T., Schram, A. (2004). How universal is behavior? A four country comparison of spite, cooperation and errors in voluntary contribution mechanisms. Public Choice, 119, 381-424.

Brosig, J., Helbach, C., Ockenfels, A., Weimann, J. (2011). Still different after all these years: Solidarity behavior in East and West Germany. Journal of Public Economics, forthcoming.

Buchan, N., Grimalda, G., Wilson, R., Brewer, M., Fatas, E., Foddy, M. (2009). Globalization and human cooperation. Proceedings of the National Academy of Sciences of the United States of America, 106, 11, 4138-4142.

Burlando, R., Guala, F. (2005). Heterogeneous agents in public goods experiments. Experimental Economics, 8, 35-54.

Burlando, R., Hey, J. (1997). Do Anglo-Saxons free-ride more? Journal of Public Economics, 64, 41-60.

Carpenter, J., Cardenas, J.C. (2011). An inter-cultural examination of cooperation in the commons.
Journal of Conflict Resolution, 55, 4, 632-651.

Cason, T., Tatsuyoshi, S., Yamato, T. (2002). Voluntary participation and spite in public good

provision experiments: an international comparison. Experimental Economics, 5, 2, 133-153.

Cederman, L., Girardin, L. (2007). Beyond fractionalization: mapping ethnicity onto nationalist insurgencies. American Political Science Review, 101, 173-185.

Chandra, L. (2004). Why ethnic parties succeed. Cambridge, Cambridge University Press.

Charles, K., Kline, P. (2006). Relational costs and the production of social capital: evidence from carpooling. Economic Journal, 116, 511, 581-604.

Easterly, W., Levine, R. (1997). Africa's growth tragedy: Policies and ethnic divisions. Quarterly Journal of Economics, 112, 4, 1203-50.

Eckel, C., Petrie, R. (2011). Face value. American Economic Review. Forthcoming.

Fafchamps, M. (2000). Ethnicity and credit in African manufacturing, Journal of Development Economics, 61, 1, 205-235.

Faillo, M., Grieco, D., and Zarri, L. (2010). Legitimate punishment, feedback, and the enforcement of cooperation. Economics Department Working paper n.16, University of Verona.

Fearon, J. (2003). Ethnic and cultural diversity by country. Journal of Economic Growth, 8,2, 195-222.

Fehr, E., Gächter, S. (2000). Cooperation and punishment in public goods experiments. American Economic Review, 90, 4, 980-994.

Ferraro, P.J., Cummings, R.G., 2007. Cultural diversity, discrimination, and economic outcomes: an experimental analysis, Economic Inquiry, 45, 2, 217-232.

Fershtman, C., Gneezy, U. (2001), Discrimination in a segmented society: an experimental approach. Quarterly Journal of Economics, 115, 351-377.

Finocchiaro Castro, M. (2008). Where are you from? Cultural differences in public good experiments. Journal of Socio-Economics, 37, 2319-2329.

Fischbacher, U., Gächter, S., Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. Economics Letters, 397-404.

Fisman, R., Miguel, E. (2007). Corruption, norms and legal enforcement: Evidence from diplomatic parking tickets. Journal of Political Economy, 115, 1020-1048.

Frank, M. G., Gilovich, T. (1988). The dark side of self and social perception: Black uniforms and aggression in professional sports. Journal of Personality and Social Psychology, 54, 74-83.

Gächter, S., Herrmann, B. (2009). Reciprocity, culture and human cooperation: previous insights and a new cross-cultural experiment, Philosophical Transactions of the Royal Society B, 364, 1518, 791-806.

Gächter, S., Herrmann, B. (2010). The limits of self-governance when cooperators get punished: experimental evidence from urban and rural Russia. European Economic Review, 55, 2, 193-210.

Goette, L., Huffman, D., Meier, S., Sutter, M. (2010). Group membership, competition, and altruistic versus antisocial punishment: evidence from randomly assigned army groups, IZA Discussion Paper N. 5189.

Greif, A. (1993). Contract enforceability and economic institutions in early trade: the Maghribi traders' coalition. American Economic Review, 83, 3, 525-548.

Guiso, L., Sapienza, P., Zingales, L. (2006). Does culture affect economic outcomes? Journal of Economic Perspectives, 20, 23-48.

Guala, F. (2011). Reciprocity: weak or strong? What punishment experiments do (and do not) demonstrate. Behavioral and Brain Sciences, forthcoming.

Habyarimana, J., Humphreys, M., Posner, D., Weinstein, J. M. (2007). Why does ethnic diversity undermine public goods provision? American Political Science Review, 101, 4, 709-725.

Henrich, J. (2004). Cultural group selection, coevolutionary processes and large-scale cooperation. Journal of Economic Behavior and Organization, 53, 1, 3-35.

Henrich, J, et al. (2006). Costly punishment across human societies. Science, 312, 1767-1770.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R. (2001). In search of homo economicus: behavioral experiments in 15 small-scale societies. American Economic Review 91, 73-78.

Herrmann, B., Thoeni, C., Gächter, S. (2008). Antisocial punishment across societies. Science, 319, 1362-1367.

Hoff, K., Kshetramade, M., Fehr, E. (2011). Caste and punishment: the legacy of caste culture in norm enforcement. Economic Journal, 121, F449-F475.

Khwaja, A. (2002). Can good projects succeed in bad communities? Collective action in the Himalayas. Unpublished manuscript, Harvard University.

Kurzban, R., Houser, D. (2005). Experiments investigating cooperative types in humans: A complement to evolutionary theory and simulations. Proceedings of the National Academy of Sciences of the United States of America, 102, 5, 1803-1807.

Levine, D. (1998). Modeling altruism and spitefulness in experiments. Review of Economic Dynamics, 1, 593-622.

Luttmer, E. (2001). Group loyalty and the taste for redistribution. Journal of Political Economy, 109, 500–528.

Mauro, P. (1995). Corruption and growth. Quarterly Journal of Economics, 110, 681–712.

Miguel, E., Gugerty, M.K. (2005). Ethnic diversity, social sanctions, and public goods in Kenya. Journal of Public Economics, 89, 2325-2368.

Ockenfels, A., Weimann, J. (1999). Types and patterns: an experimental East-West-German comparison of cooperation and solidarity. Journal of Public Economics, 71, 275-287.

Offerman, T., Sonnemans, J., Schram, A. (1996). Value orientation expectations, and voluntary contributions in public goods. Economic Journal, 106, 817-845.

Ones, U., Putterman, L. (2007). The ecology of collective action: A public goods and sanctions experiment with controlled group formation. Journal of Economic Behavior and Organization, 62, 495-521.

Rand, D.G., Armao IV, J.J., Nakamaru, M., Ohtsuki, H. (2010). Anti-social punishment can prevent the co-evolution of punishment and cooperation. Journal of Theoretical Biology, 265, 624-632.

Roth, A.E., Prasnikar, V., Okuno-Fujiwara, M., Zamir, S. (1991). Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An experimental study. American Economic Review, 81, 5, 1068-1095.

Rustagi, D., Engel, S., Kosfeld, M. (2010). Conditional cooperation and costly monitoring explain success in forest commons management. Science, 330, 961-965.

Schelling, T. (1971). On the ecology of micromotives. The Public Interest, 25, 61-98.

Sober, E. ,Wilson, D. (1998). Unto others: the evolution and psychology of unselfish behavior. Harvard University Press.

van Dijk, F., Sonnemans, J., van Winden, F. (2002). Social ties in a public good experiment. Journal of Public Economics, 85, 275-299.

Vigdor, J. (2002). Interpreting ethnic fragmentation effects. Economics Letters, 75, 271-276.

Vigdor, J. (2004). Community composition and collective action: analyzing initial mail response to the 2000 census. Review of Economics and Statistics, 86, 1, 303-312.

Weimann, J. (1994). Individual behavior in a free riding experiment. Journal of Public Economics, 54, 185-200.

Zizzo, D. (2010). Experimenter demand effects in economic experiments. Experimental Economics, 13, 1, 75-98.

Figures and tables

Table 1: Cost of punishment points given

Points	0	1	2	3	4	5	6	7	8	9	10
Costs	0	1	2	4	6	9	12	16	20	25	30

Table 2. Subjects' distribution by country

Session	Treatment	IT	AL	BR	CAM	CD	ER	ETH	GH	LT	MA	MD	MK	MOC	SRB	UA	Tot non Italian	Tot
1	Baseline	17							1						1	1	3	20
2	Baseline	12			2	1	1			2		1	1				8	20
3	Baseline	13	2	1													3	16
4	Restricted	11	2		1						1		1				5	16
5	Restricted	18						1				1					2	20
6	Restricted	18	2														2	20
	Tot	89	6	1	3	1	1	1	1	2	1	2	2	0	1	1	23	112

Table 3. Average contribution levels by session

Treatment	Tot non Italian	(1) Average contribution Italians first period	(2) Average contribution foreigners first period	(3) Average contribution all subjects all periods	(4) Average contribution italian all periods	(5) Average contribution foreigners all periods
Baseline	3	6.64	5	12.25	12.72	9.58
Baseline	8	7.83	2.75	3.84	4.94	2.20
Baseline	3	6	6.66	9.03	9.14	8.53
Full.R	5	8.45	4	11.79	14.41	5.98
Full.R	2	11.38	3	15.5	16.44	10.06
Full.R	2	10.11	12	15.74	15.57	17.25
Tot.	23					

Figure 1. Average contribution in the Baseline Treatment: natives vs. foreigners

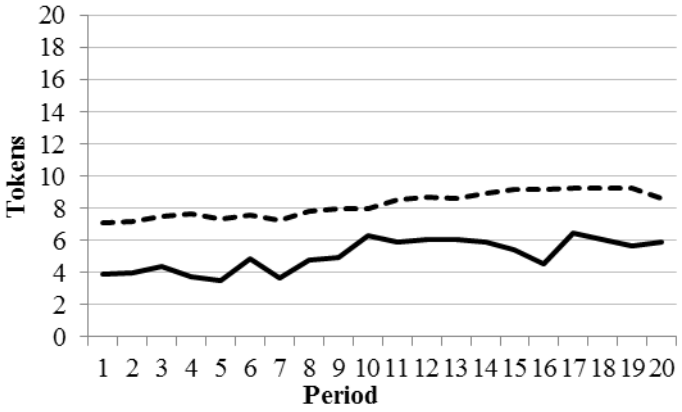

Figure 2. Average contribution level in the Restricted Treatment: natives vs. foreigners

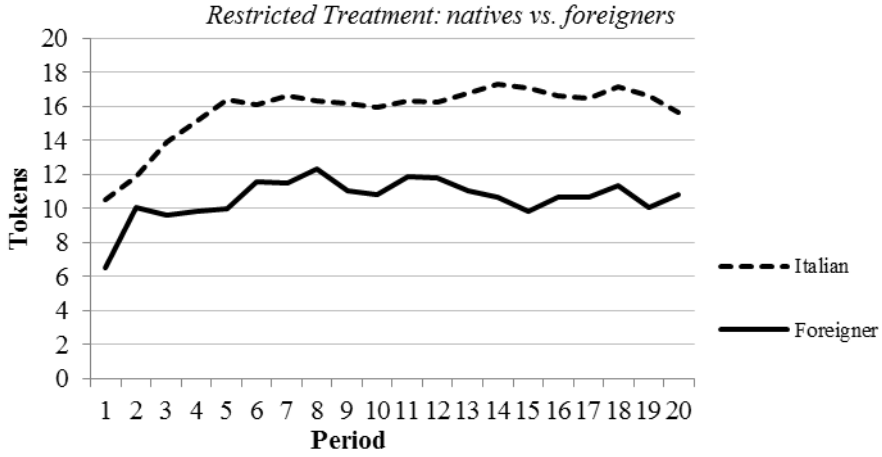

Table 4. Determinants of contribution levels

Contribution at t	**Random effect Tobit**
Gender	-1.80
	(1.21)
Age	0.55 ***
	(0.21)
Restricted	7.60 ***
	(2.62)
Nationality	5.62 ***

	(2.06)
Period 1-7	-2.11 ***
	(0.41)
Period 8-15	-0.57
	(0.40)
Restricted*Period 1-7	-2.31***
	(0.61)
Restricted*Period 8-15	0.21
	(0.61)
N. Experiments	0.00
	(0.11)
Major	3.05 **
	(1.34)
Nationality*Restricted	3.85
	(2.87)
Constant	-9.41 *
	(5.69)

Number of obs = 3040; Wald chi2(11) = 274.03; Log likelihood = -6799.0776

Dependent variable: Contribution, which takes values from 0 to 20;
Gender: dummy variable which takes value 1 if the subject if a female;
Age : age of the subject;
Restricted: dummy variable which takes value 1 in the Restricted treatment.
Nationality: dummy variable which takes value 1 if the subject is Italian.
Period 1-7: dummy variable which takes value 1 if the period is between 1-7;
Period 8-15: dummy variable which takes value 1 if the period is between 8-15;
N. Experiments: number of previous experiments the subject has undertaken.
Major = dummy variable which take value 1 if the subject is an Economics major.

* Significant a 10%.
** Significant at 5%.
*** Significant at 1%.

Table 5. The effect of antisocial punishment on contribution level

Contribution	**Random effect Tobit**
Antisocial received lag	-0.96**
	(0.41)
Gender	0.61
	(1.65)

Age	0.77**
	(0.33)
Nationality	5.56***
	(1.95)
N. Experiments	0.15
	(0.17)
Major	2.72
	(1.81)
Period 1-7	-2.46***
	(0.46)
Period 8-15	-0.57
	(0.35)
Period 1 -5*Antisocial received lag	1.47***
	(0.46)
Period 8 -15*Antisocial received lag	0.00
	(0.50)
Constant	-16.00*
	(8.54)

Number of obs= 1444; Wald chi2(6) =9.81; Log likelihood = -3495.063

Dependent variable: Contribution, which takes values from 0 to 20;
Antisocial received lag: number of antisocial punishment point received in the previous period
Gender: dummy variable which takes value 1 if the subject if a female;
Age : age of the subject;
Nationality: dummy variable which takes value 1 if the subject is Italian.
N. Experiments: number of previous experiments the subject has undertaken.
Major = dummy variable which take value 1 if the subject is an Economics major.
Period 1-7: dummy variable which takes value 1 if the period is between 1-7;
Period 8-15: dummy variable which takes value 1 if the period is between 8-15;

Table 6. Antisocial punishment and its determinants.

Percentage of anti-social points received	**Random effect Tobit**
Contribution	-0.45***
	(0.07)
Age	0.09
	(0.14)
Gender	2.92***

	(0.33)
Nationality	-0.35
	(0.82)
N. Experiments	0.88
	(0.09)
Major	-0.17
	(0.86)
Constant	1.65
	(3.83)

Number of obs = 534; Wald chi2(8) = 39.39; Log likelihood = -233.3322

Dependent variable:
Perc. Antisocial received = (antisocial punishment points received /total amount of punishment points received) x100

Contribution: contribution
Gender: dummy variable which takes value 1 if the subject if a female;
Age : age of the subject;
Nationality: dummy variable which takes value 1 if the subject is Italian.
N. Experiments: number of previous experiments the subject has undertaken.
Major = dummy variable which take value 1 if the subject is an Economics major.

* Significant a 10%.
** Significant at 5%.
*** Significant at 1%.
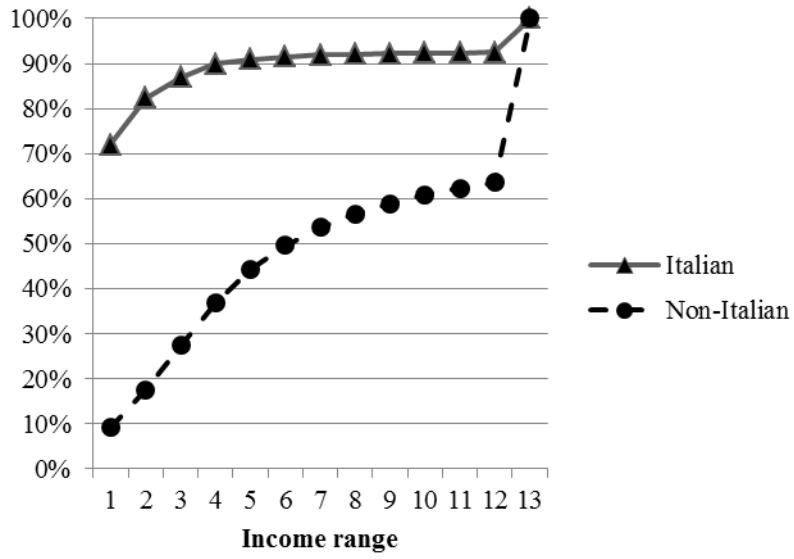
Figure 3. Income of University of Trento students (cumulative distribution) (n=11,513)