



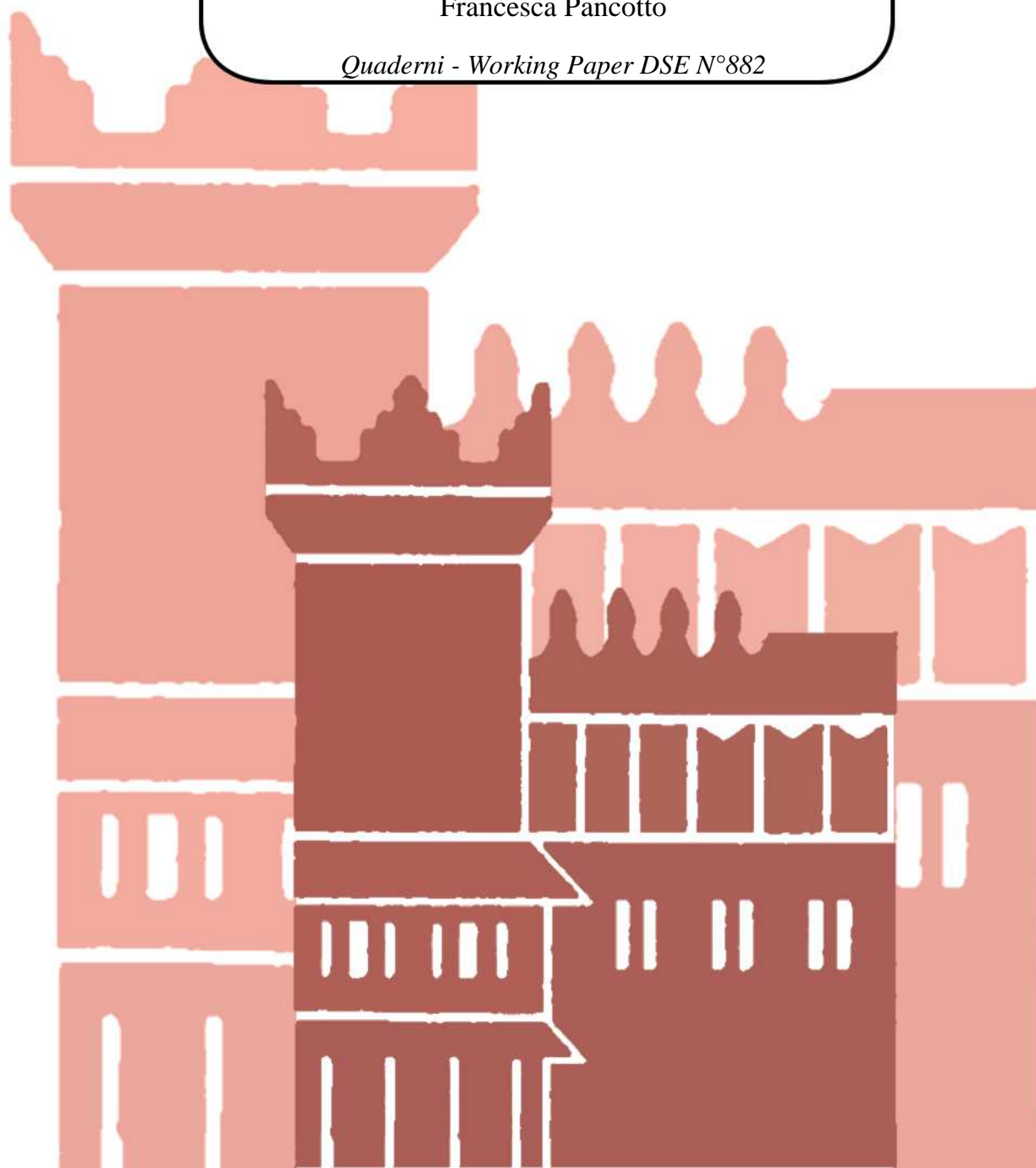
ISSN 2282-6483

Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

**Cooperation Hidden Frontiers:
The Behavioral Foundations of the
Italian North-South Divide**

Maria Bigoni
Stefania Bortolotti
Marco Casari
Diego Gambetta
Francesca Pancotto

Quaderni - Working Paper DSE N°882



Cooperation Hidden Frontiers: The Behavioral Foundations of the Italian North-South Divide

Maria Bigoni* Stefania Bortolotti† Marco Casari‡
University of Bologna

Diego Gambetta§
European University Institute & Oxford University

Francesca Pancotto¶
University of Modena and Reggio Emilia

November 8, 2013

Abstract

Socio-economic performance differs not only across countries but within countries too and can persist even after religion, language, and formal institutions are long shared. One interpretation of these disparities is that successful regions are characterized by higher levels of trust, and, more generally, of cooperation. Here we study a classic case of within-country disparities, the Italian North-South divide, to find out whether people exhibit geographically distinct abilities to cooperate independently of many other factors and whence these differences emerge. Through an experiment in four Italian cities, we study the behavior of a sample of the general population toward trust and contributions to the common good. We find that trust and contributions vary in unison, and diminish moving from North to South. This regional gap cannot be attributed to payoffs from cooperation or to institutions, formal or informal, that may vary across Italy, as the experimental methodology silences their impact. The gap is also independent of risk and other-regarding preferences which we measure experimentally, suggesting that the lower ability to cooperate we find in the South is not due to individual “moral” flaws. The gap could originate from emergent collective properties, such as different social norms and the expectations they engender. The absence of convergence in behavior during the last 150 years, since Italy was unified, further suggests that these norms can persist overtime. Using a millennium-long dataset, we explore whether the quality of past political institutions and the frequency of wars could explain the emergence of these differences in norms.

Keywords: Trust, social norms, experiments, long-term persistence, conflicts, institutions.

JEL codes: C90, D03

*Bigoni: Department of Economics, University of Bologna, Piazza Scaravilli 2, 40126 Bologna, Italy; maria.bigoni@unibo.it, Phone: +39 051 209 8122, Fax: +39 051 0544 522.

†Bortolotti: Department of Economics, University of Bologna, Piazza Scaravilli 2, 40126 Bologna, Italy; stefania.bortolotti@unibo.it, Phone: +39 051 209 8135.

‡Casari (corresponding author): Department of Economics, University of Bologna, Piazza Scaravilli 2, 40126 Bologna, Italy; marco.casari@unibo.it, Phone: +39 051 209 8662, Fax: +39051 209 8493.

§Gambetta: European University Institute, Badia Fiesolana - Via dei Roccettini 9, I-50014 San Domenico di Fiesole, Italy, diego.gambetta@eui.eu.

¶Pancotto: Department of Communication and Economics, University of Modena and Reggio Emilia, Viale A. Allegri 9, 42121 Reggio Emilia (RE) – Italy, francesca.pancotto@unimore.it, Phone: +39 0522 523 264.

1 Introduction

A political border can make all the difference for the socio-economic outcomes of two otherwise identical regions. Nogales, for instance, a city cut in half by the US-Mexican border, is a vivid illustration of the powerful driving force of formal institutions (Acemoglu and Robinson, 2012). In other cases, neighboring regions under the same formal institutions show such sharp discontinuities as if they were divided by an invisible frontier. Two villages in the Alps, Tret and St. Felix, for example, have sat side by side for centuries in the same jurisdiction and under identical environmental conditions, and yet exhibit dramatic and persistent differences in socio-economic patterns (Cole and Wolf, 1974). The reason seems to lie in the entrenched but distinct preferences and social norms.

Here we study a classic grand case of the latter type, Italy, whose regions present dramatic differences in terms of economic success. Since its Unification in 1861, the country's formal institutions are shared, yet the disparities between the Northern and the Southern regions have persisted along many social and economic dimensions (Putnam et al., 1993; Guiso et al., 2004; Daniele and Malanima, 2011). These disparities have received much scholarly attention over the last fifty years. To our knowledge, however, this is the first time that a controlled field experiment was employed to uncover their behavioral foundations in terms of the ability to cooperate.

Cooperation – which we define as a joint effort which benefits the group at some cost to the individual – has been identified as a major determinant of economic success (Nowak and Highfield, 2011), but its variability across regions can be difficult to measure; the reason is that observed differences often derive not only from disparities in terms of formal and informal institutions but also from differences in the payoffs from cooperation. By contrast, the employed experimental methodology removes the confounding factors that could shape behavior such as the degree of corruption, the varying efficiency of the public administration or the level of mafia intrusion. Should regional disparities emerge as distinct behavioral responses to otherwise identical experimental conditions, the implication would be that they are not so much due to differences in institutions or incentives but rather to differences in preferences or in social norms.

Italy offers the necessary variability in socio-economic performances and local histories to investigate cooperative behavior, while, at the same time, exhibiting constancy in several relevant factors such as institutions, language and dominant religion. These characteristics allow us to go beyond previous experiments based on cross-country evidence, which have shown that cooperation co-varies with factors shared at society level, such as market integration, production technology, religion, the quality of the rule of law, and the degree of globalization (Henrich et al., 2010; Herrmann et al., 2008; Buchan et al., 2009).

In a controlled experiment, we study whether differences in cooperation rely on differences in individual traits or are, by contrast, the result of emergent properties of societies. Edward

Banfield famously claimed that the origin of the North-South gap in Italy lies in “moral” flaws of Southerners, whose only concern would be with the welfare of self and of immediate family members with utter disregard for anyone else (Banfield and Fasano, 1958). Alternatively, the behavioral gap could be due to collective dimensions, such as social norms, group cohesion, or mutual expectations. We performed an in-depth analysis of each society, by collecting multiple measures of in-group cooperation and of individual traits, in an experiment with a large pool of participants (about 150 per city). The measures of cooperation were obtained through a three-player trust game and a public good game; the measures of individual traits were obtained through a set of dictator games and a lottery. All these dimensions of behavior, which are rarely studied jointly in cross-cultural studies, can in theory be expected to affect cooperation. In addition, we combine the experimental findings with a historical record about the quality of institutions and the frequency of external violent conflicts during the last millennium. On this basis we construct a ranking of plausibility among the various conjectures present in the literature on how collective experiences in the distant past shape current behavior (Tabellini, 2008; Benhabib et al., 2011; Nunn and Wantchekon, 2011).

The experiment was conducted in four Italian cities, with a representative sample of the general population (N=618). We chose two cities in the North and two in the South, and in order to uncover deep-rooted dispositions, we focused on medium-size cities large enough to study cooperation with individuals beyond the immediate family circle, but small enough to have a stable community, in which dispositions are well known and shared. For each city the recruited sample was balanced in terms of age, sex, and occupation, which is uncommon in previous experimental cross-country studies. The representativeness of the sample is particularly relevant for a research on the sources of cooperation: first, because it increases the comparability across societies; second, because it boosts the external validity of the results when compared with studies run with college students. In order to capture local emergent properties, it is essential for all participants to belong to the same community and to experience repeated interactions. For this purpose, we chose participants who were born and resided in the county where we ran the experiments, an information which we divulged to participants.

We report four main findings. First, trust and contributions to the public good vary in unison. Second, we find that both trust and contributions are higher in the North than in the South. Third, differences in risk and other-regarding preferences cannot account for the observed gap in cooperation, hence suggesting that the divide is an emergent property of societies and is not due to individual “moral” flaws. Finally, we find that the quality of past political institutions and the frequency of external violent conflicts in the last millennium could explain the emergence of these differences in norms.

The remainder of the paper is organized as follows. Section 2 describes the the experimental design and the procedures used to select the sample and the experimental locations. Section

3 presents the main experimental findings and Section 4 discusses the historical data and their link to present measures of cooperation. Section 5 concludes.

2 Sample and design

2.1 Experimental locations and sample selection

We selected the experimental locations following rigorous criteria and with three goals in mind: *(i)* to cover Italian counties (*‘Province’* in Italian)¹ both in the North and in the South macro-areas; *(ii)* to maximize the difference relative to social capital within each macro-area; and *(iii)* to have medium-size cities. First, in order to cover both macro-areas, we included two counties in the North and two in the South of Italy.² We excluded the Sardegna Island because of its geographical and historical peculiarities.

The second selection criterion pertained the level of social capital in each county. The aim was to increase the likelihood of sampling counties that would differ in terms of trusting and cooperative behavior. Hence, the choice of the counties was made so to maximize the difference in their social capital level. The proxies for social capital were those widely adopted in the literature (i.e., Putnam et al., 1993; Guiso et al., 2004, 2006): association density,³ electoral participation,⁴ and blood donations.⁵

To aggregate these multiple measures of social capital into a single index, we adopted the following procedure. First, we separately ranked all Italian counties according to each of the three dimensions. In each ranking, position 1 was assigned to the county with the highest value in the current dimension and the last position to the county with the lowest value. The

¹As of December 2000, there were 103 counties in Italy; other 7 counties were introduced by the Italian Government between 2001 and 2011, which we did not include due to lack of data for one or more of the adopted selection criteria.

²North and South are defined according to the official classification of the National Bureau of Statistics, Istat, see <http://www.istat.it/it/archivio/6789>. The regions of the North are Piemonte, Valle d’Aosta, Lombardia, Trentino-Alto Adige, Veneto, Friuli-Venezia Giulia, Liguria, Emilia-Romagna. The regions of the center are: Toscana, Umbria, Marche, Lazio. The regions of the South are Abruzzo, Molise, Campania, Puglia, Calabria, Sicilia, and Sardegna.

³ Association density at the county level in year 2000 is computed as the number of associations (cultural, leisure, artistic, sports, environmental, and any kind of nonprofit associations) per 100,000 inhabitants (Source: ISTAT). We considered registered voluntary associations according to law 291/91 per 100,000 inhabitants, collected in *‘Primo censimento istituzioni nonprofit in Italia, 1999.’* We excluded professional and religious associations in accordance to the convention adopted in the literature on social capital (Putnam et al., 1993; La Porta et al., 1997).

⁴ Electoral participation considers the average per county turnout, expressed as the percentage of eligible voters in all referenda held in Italy from 1946 to 1999 (Guiso et al., 2004 on original data of the Italian Ministry of Internal Affairs). The counties of Belluno and Aosta have missing values for referenda turnout. We assigned to Belluno the average value for the Veneto region and to Aosta the average value of the Piemonte region. Referenda are chosen in place of political elections as they are considered a better proxy of the desire of civic participation, rather than the regular Parliamentary elections. The reasons are varied: voting at referenda is not mandatory in Italy, the issues on the ballot in referenda are less related to local interests, and referenda are immune from possible contamination from the so called ‘exchange vote.’ The exchange vote is an illegal practice according to which people may receive a payment in order to cast a vote for a particular candidate (See also Guiso et al., 2010; Putnam et al., 1993).

⁵ We considered the number of blood donations per 1,000 inhabitants in each county in year 2002 (Cartocci, 2007). Data for Lazio and Puglia are available at the regional rather than county level and data for the counties of Gorizia and Trieste are the average value for the two counties.

county with the highest level of social capital was picked according to a minimax criterion: the score for each county was computed by taking the maximum among the three rankings and then selecting the county with the lowest score. The county with the lowest level of social capital was picked according to a maximin criterion: the score for each county was computed by taking the minimum value among the values of three rankings and then selecting the county with the highest score. The procedure was performed including all Italian counties, divided into two groups: North and South. The Central counties are placed in one of the two clusters.⁶ The county with the highest (lowest) social capital in the North was Ravenna (Cuneo), while the county with the highest (lowest) social capital in the South was Ragusa (Crotone). The procedure is robust to excluding the counties of Central Italy.

Finally, we considered the size of the city where the experiment would take place. Although participants were recruited from the whole county, we aimed at cities that could be compared in terms of size (between 50,000-100,000 inhabitants). Consider that the average size of an Italian county capital is 160,428 inhabitants, a value that lowers to 94,824 when excluding the five largest Italian cities (Roma, Milano, Torino, Palermo, Napoli). The corresponding median populations are 82,367 and 72,329, respectively. The average municipality in Italy is much smaller: 7,492 inhabitants.⁷ We wanted to avoid both extremely small towns and large metropolitan cities to have a measure of deep-seated social norms specific to the area. As we aimed at studying interactions among strangers, very small cities could make difficult to achieve the required anonymity among participants. On the other hand, in large cities the population is more diverse and it has higher mobility. After having identified the counties of interest, we restricted the attention to medium-size cities. All the cities are provincial capitals except for Faenza. In particular, Faenza (56,992 inhabitants) was selected instead of Ravenna as the latter has more than 100,000 inhabitants (155,997 at 2009).⁸ Thus, the four cities selected are: Cuneo, Faenza, Crotone, and Ragusa (see Table A-1 in Appendix).

The four Italian cities chosen as experimental locations are geographically distant both in terms of kilometers and traveling times. Consider traveling by car from the northernmost town of Cuneo to the southernmost town of Ragusa. According to Google maps, it takes 4 hours and 25 minutes to drive from Cuneo to Faenza (458 km), and then an additional 9 hours and 46 minutes to reach Crotone (919 km). From Crotone it takes 6 hours and 13 minutes to arrive in Ragusa (425 km).

Study participants were recruited among the general population of each county. To be eligible for the study, subjects had to: *(i)* be at least 18 years old; *(ii)* be born in the county; *(iii)* be resident in the county; *(iv)* have a good knowledge of spoken and written Italian. The purpose of these requirements was to ensure that subjects shared the social norms of the

⁶Marche, Umbria and Toscana to the Northern group and Lazio to the Southern group. Latitude was used to decide the assignment.

⁷Source: Istat, <http://demo.istat.it/pop2010/index.html>.

⁸Source: Istat: <http://demo.istat.it/pop2010/index.html>

place or had at least a profound knowledge of them. These restrictions were explained during the recruitment process and the experimenters double checked this condition by looking at the ID of each participant before the experimental sessions. At the beginning of each session, the experimenter made it public that all subjects in the room were born and resident in the county (or at least in the region) in order to make this information common knowledge. Among the 581 subjects for which we have questionnaire data, 97.3% were born in the county, 2.4% in the region, and 0.3% outside the region. About 92.7% of the participants turned out to be at least second generation natives of the county, based on the reported birthplace of their mother and father (i.e., at least one of their parents was born in the county).

We wanted a representative sample of the Italian population in respect to age, sex, and employment status, since these demographics characteristics could be important to identify social norms. The sample was stratified according to three categories of age (18-39, 40-59, 60 and older), two of sex (male and female), and three of employment status (employed; housewives and retired; others, including students and unemployed). For the composition of the target sample, we referred to the 2009 statistics on the Italian population.⁹ Table A-2 in Appendix summarizes the socio-demographic characteristics of the actual sample. Two professional companies –Metis-Ricerche and Demoskopea– were hired for the recruitment of subjects.¹⁰

2.2 Experimental design and procedures

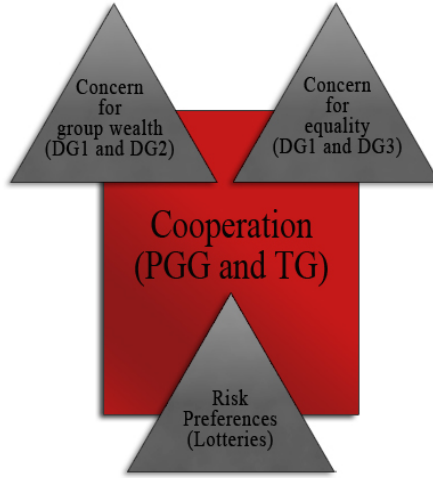
The study included two types of sessions: roughly half of the subjects participated in the Public Goods (PGG) session and half in the Trust Game (TG) session along with three modified Dictator Games. Everyone undertook a choice over lotteries. The experiment comprised collective tasks (square in Figure 1) and individual tasks (triangles in Figure 1).

Each PGG session comprised five tasks presented in a fixed order. In the first task (Lottery), we employed a procedure similar to Eckel and Grossman (2008) to elicit risk preferences. Subjects had to choose from a list of six lotteries; each lottery had two possible outcomes (high, low) that occurred with equal probability (Table 1). Only two randomly selected subjects per session received a payment for the lottery task. To avoid any possible carry over effect, lotteries earnings and winners were determined only at the end of the session, by a manual draw of a colored ball out of a bag. In the second task (PGG-Standard), subjects faced 8 rounds of a standard Public Goods game. In each round, subjects were randomly assigned to groups of $N = 4$ according to a strangers matching protocol and they could contribute either 0, 6, 14 or 20 money units (MUs) to a group project. Every MU invested in the project was doubled and shared equally among group members, hence generating 0.5

⁹Inhabitants at 1st of January 2009. Age range: 18-39 years, 34.8%; 40-59 years, 34.6%; 60 and more, 31.6%. Sex: male, 48%; female 52%. Employment status: employed, 42%; housewives and retired, 37%; others 21%. Source: <http://demo.istat.it/pop2009/index1.html>

¹⁰ See the Appendix for further details on the recruitment process.

Figure 1: Experimental Tasks



MUs for everyone in the group. Tasks three, four, and five included two variations of the PGG and are not included in the present paper. Task three (PGG-Punishment) consisted of 8 rounds of a PGG with punishment.¹¹ Task four was identical to task two but lasted 4 rounds. Finally, in task five subjects faced 8 rounds of a public good with a threshold (PGG-Threshold).¹²

Table 1: Lottery Task

| | High payoff (orange ball) | Low payoff (white ball) | Risk-aversion (CRRA) |
|-----------|------------------------------|----------------------------|-------------------------|
| Lottery 1 | 17.5 | 17.5 | >3.64 |
| Lottery 2 | 22.5 | 15.0 | 3.46-1.16 |
| Lottery 3 | 27.5 | 12.5 | 1.16-0.70 |
| Lottery 4 | 32.5 | 10.0 | 0.70-0.50 |
| Lottery 5 | 37.5 | 7.5 | 0.50-0 |
| Lottery 6 | 44.0 | 1.0 | <0 |

Notes: As an index of risk-aversion we take the ranges of the coefficient of constant relative risk aversion (CRRA) implied by subjects' choices. These indexes were not shown to subjects. Payoffs are in Euros.

Each TG session included four tasks that were presented in a fixed order. The first task was a Lottery and it was identical to the first task of the PGG sessions. In the second task, subjects faced three modified Dictator Games (DGs): DG1, DG2, and DG3. Session participants were divided into groups of three members. In each game, a subject had to choose how to allocate amounts of MUs among himself and the other two group members. Table 2 presents the six alternative allocations available in each situation. Each subject was asked to play as the dictator (red player) in DG1, DG2, and DG3; at the end of the session,

¹¹In each round, after playing the standard version of the PGG, subjects could assign 0, 1 or 2 deduction points to each group member. Each deduction point had a cost of 1 MU for the punisher and reduced the earnings of the targeted subject by 4 MUs. Punishment decisions were simultaneous and new groups were formed at the end of each period.

¹²In this case the public good was provided only if all four subjects in the group contributed 20; in case the threshold was not reached, contributions were not refunded. Data for this task are not analyzed in the current paper.

only one of the games was selected at random for payment. Importantly, roles in the game were then randomly assigned to determine earnings. We now describe each dictator game.

In DG1 the dictator (red) always earned 160 MUs and he faced choices between equality vs. group wealth. While allocation 1 ensured equal earnings to all three group members (E), allocation 6 delivered the highest sum of earnings for the group (W). Allocations 2, 3, 4, and 5 provided intermediate situations between E and W. A merely self-interested dictator would be indifferent among all the available allocations, while the choice of a dictator who cares for others' welfare would reveal his relative preferences for equality over group wealth. In DG2 the dictator faced a tradeoff between self-interest and group wealth. As a matter of fact, DG1 and DG2 shared two main features. First, allocation 6 (W) was identical in both games. Second, group wealth varied with the allocation $k=1,\dots,6$ but was identical across DG1 and DG2 (Table 2). The main difference between DG1 and DG2 resided in the earnings of the dictators; while dictator's earnings are constant in the former situation, they vary from 160 to 190 MUs in the latter. Allocations 1 through 6 in DG2 were designed to have the same level of inequality as measured according to Fehr and Schmidt (1999)'s model, under the assumption of equal weights for disadvantageous and advantageous inequality.¹³ Hence, a purely self-interested dictator would always choose allocation 1 (S) over all other allocations in DG2. In contrast, in DG3 group wealth is kept constant (480 MUs) in all six allocations and the dictator faces a tradeoff between self-interest and equality. Allocation 1 (S) yields the highest earnings to the dictator but the distribution is highly unequal, whereas allocation 6 (E) presents equality of earnings and yields the lowest earnings for the dictator. A self-interested dictator will choose allocation 1 and earn 190 MUs, while a dictator concerned with inequality may choose allocation 6 and earn 160 MUs.

In the third task (TG-Baseline), subjects were divided into groups of three and played 10 rounds of a trust game. The truster decided first and could trust or not two trustees. If the truster did not trust, everyone in the group earned 20 (MUs) and the trustees had no choice to make. If the truster trusted and at least one trustee reciprocated, the truster earned 36 MUs, and each trustee earned 30 MUs. When no trustee reciprocated, the truster earned 0 MUs and each trustee earned 48 MUs. At the end of each round, roles were randomly re-assigned and new groups were formed according to a strangers protocol (in each session there were two matching groups). Finally, the fourth task consisted of 10 additional rounds of a modified TG, which are not reported in the present paper.¹⁴

All sessions were held in hotel conference rooms or educational centers located near to the city center and each location was devoid of any political or religious connotation (see Figure A-1 in the Appendix). They were easy to reach for the participants, i.e. accessible by car and by public transport, and near or at well-known locations. We moved the BLESS laboratory

¹³The difference in inequality measured according to Bolton and Ockenfels (2000)'s model is also minimal in these allocations.

¹⁴A trustee who does not reciprocate while the other does earned 20 (instead of 30). For a more detailed description of the game and the equilibrium structure, see the Coordination version of the Collective Trust Game discussed in Bigoni et al. (2013).

Table 2: Modified 3-player Dictator Games

| DG1: Equality vs. Group Wealth | | | | | | |
|--------------------------------|------------|------------|------------|------------|------------|------------|
| Allocations | | | | | | |
| | 1 (E) | 2 | 3 | 4 | 5 | 6 (W) |
| You (red) | 160 | 160 | 160 | 160 | 160 | 160 |
| Other participant (black) | 160 | 154 | 148 | 142 | 136 | 130 |
| Other participant (white) | 160 | 196 | 232 | 268 | 304 | 340 |
| <i>Group wealth</i> | <i>480</i> | <i>510</i> | <i>540</i> | <i>570</i> | <i>600</i> | <i>630</i> |

| DG2: Self-Interest vs. Group Wealth | | | | | | |
|-------------------------------------|------------|------------|------------|------------|------------|------------|
| Allocations | | | | | | |
| | 1 (S) | 2 | 3 | 4 | 5 | 6 (W) |
| You (red) | 190 | 184 | 178 | 172 | 166 | 160 |
| Other participant (black) | 40 | 58 | 76 | 94 | 112 | 130 |
| Other participant (white) | 250 | 268 | 286 | 304 | 322 | 340 |
| <i>Group wealth</i> | <i>480</i> | <i>510</i> | <i>540</i> | <i>570</i> | <i>600</i> | <i>630</i> |

| DG3: Self-Interest vs. Equality | | | | | | |
|---------------------------------|------------|------------|------------|------------|------------|------------|
| Allocations | | | | | | |
| | 1 (S) | 2 | 3 | 4 | 5 | 6 (E) |
| You (red) | 190 | 184 | 178 | 172 | 166 | 160 |
| Other participant (black) | 40 | 64 | 88 | 112 | 136 | 160 |
| Other participant (white) | 250 | 232 | 214 | 196 | 178 | 160 |
| <i>Group wealth</i> | <i>480</i> | <i>480</i> | <i>480</i> | <i>480</i> | <i>480</i> | <i>480</i> |

from Bologna to the locations of the experiment, by using a van. The laboratory hardware and set-up were identical across all cities – 32 notebook computers connected through a wireless network to a laptop server. Upon arrival, subjects were seated at a desk; visual contact among participants was not possible, and no form of communication was allowed during the experiment. All participants signed a consent form and a data release form. At the beginning of each task, the relevant instructions were distributed and read out loud. The experimenter who read the instructions was the same in all sessions and in all cities. Before each task, subjects had to answer a computerized quiz to ensure correct understanding of the game before each task, with the exception of the lottery. Our subject pool included a rather large number of elderly and uneducated people: this called for a user friendly interface and simplified tasks. We largely relied on graphical elements to make the task more intuitive;¹⁵ choices were made by simply touching the screen. Indeed, there was no need to type on the keyboard or use the mouse. In an effort to reduce the complexity of the decision tasks, we limited the number of available options in each decision. At the end of the session, we administered a computerized questionnaire including several measures of risk, trust, trustworthiness, and social capital, in addition to socio-demographic information.

A total of 618 subjects participated in the study, which was conducted between March and October 2011. The number of participants in a session ranged between 24 and 32, and everyone participated in only one session. Sessions were run in the evening or on Saturday in order to favor a wider participation. The experiment was programmed with z-tree

¹⁵In programming our interfaces, we took inspiration from the first wave of experiments conducted at the Internet Laboratory for Experimental Economics, iLEE (for further details see: <http://www.econ.ku.dk/cee/ilee/description/ilee1/>).

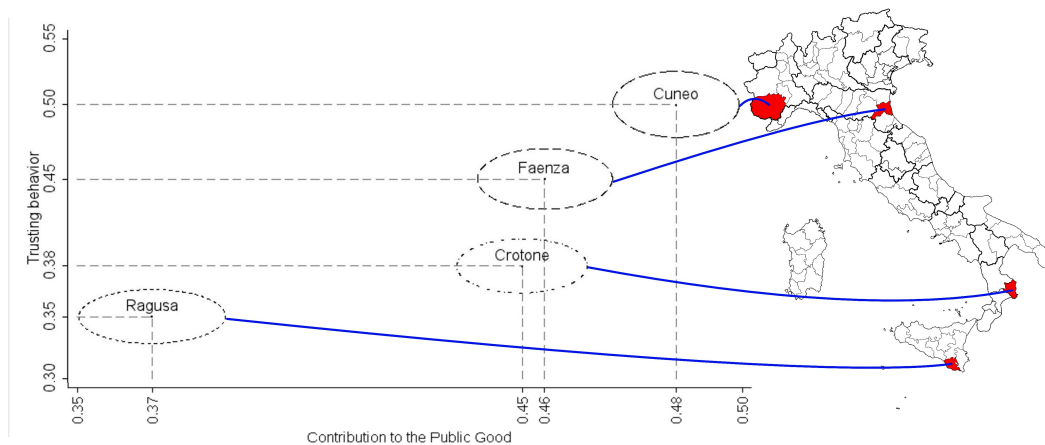
(Fischbacher, 2007) and conducted using the mobile Bologna Laboratory for Experiments in Social Sciences (BLESS). The average PGG (TG) session lasted about 2 hours (1 hour and 30 minutes). Subjects received 30 Euros gasoline vouchers for showing up on time, plus a cash payment corresponding to the sum of their earnings in each part. Payments were made privately right after the end of the experiment; average per-capita earnings were 16.5 (17.5) Euros in cash in PGG (TG) sessions plus 30 Euros in gasoline vouchers.¹⁶

3 Results

3.1 Voluntary contributions and trust go hand in hand

We found that the higher is trust in the TG the higher are the levels of voluntary contribution in the PGG. Figure 2 reports the average contribution in the PGG (as a fraction of the endowment) and the average trust frequency in the TG across all rounds. The dashed ellipses are drawn at a 95% confidence level after a bootstrapping procedure: for every county, we randomly draw, with repetition, 10,000 pairs of observations from our sample. TG and PGG involved different participants – which makes the measures fully independent of each other – and tasks with different parameters and types of strategic interaction. Yet, results are agreeing, which indicates the presence of a robust behavioral trait. On both tasks, the most cooperative city is Cuneo (0.502 in TG and 0.485 in PGG), and the least cooperative is Ragusa (0.346 in TG and 0.365 in PGG).

Figure 2: Cooperation across Italy



The ranking of cities according to average trust is identical to the ranking based on contributions. This novel evidence that trust and contributions vary in unison across cities suggests that they share a common root, which we identify as the collective ability to cooperate. Further support to this result is provided by regression analysis (Table 3). Since trust

¹⁶Reported payments include also earnings for the fourth task of the TG and fifth task of PGG(PGG-threshold).

and contributions are elicited on two different pools of subjects, it is not possible to compare them directly, at the individual level. We thus used a linear regression, where the dependent variable is the average individual contribution in the PGG, and the only independent variable in Model 1 is the frequency of trust measured at the county level (Table 3). To account for possible session effects, in this and in all the following regressions standard errors are robust for clustering at the session level. Results indicate that an increase by 10 percentage points in the frequency of trust corresponds to a highly significant increase of the average contribution, by about 1.5 points.

In a second specification (Table 3, Model 2), we include additional regressors to control for individual characteristics, elicited in the final questionnaire. More specifically, Model 2 includes a dummy taking value 1 for males (*Male*), two dummies to control for age (*Age 40-59* and *Age 60 and above*), two dummies to control for the occupational status of the subject (*Housewife/retired* and *Student/Unemployed*), and one dummy taking value one for those subjects that had troubles answering the control questions (*Low understanding*).¹⁷ Results from Model 2 indicate that the positive and significant relation between trust and individual contributions in the PGG is not driven uniquely by differences in the individual characteristics of our subjects across cities.¹⁸

3.2 Cooperation levels display a North-South ordering

In line with the theories that link cooperation and economic development (Zak and Knack, 2001; Knack and Keefer, 1997), we find that cooperation is higher in the North and lower in the South. Contributions in the PGG across cities are ordered along a North-South continuum (Jonckheere-Terpstra test: p-value= 0.037, z=2.09, n=13, two-sided).¹⁹ Similarly, the frequency of trusting behavior also displays a North-South ordering (Jonckheere-Terpstra test: p-value=0.009, z=- 2.61, n=18, 2 independent observations per session, two-sided).

Figure 2 and non-parametric tests suggest that not only trust and contributions in the PGG are strongly correlated, but they are also aligned with cities' latitude. Further support to this result is provided by a series of linear regressions, where the dependent variable is either the average contribution to the PGG (Table 4) or the frequency of trustful choices in the TG (Table 5) – both taken at the individual level – and we include each city's latitude as a regressor (Model 1). In Model 2 we add controls for the same individual characteristics

¹⁷To account both for subjects that were particularly slow in answering the control questions, and for those who made several mistakes, we attribute value 1 to the dummy variable *Low understanding* for all subjects who were in the last decile either according to their total answering time, or according to their total number of mistakes. This measure is built at the city level and separating the two treatments. Due to technical problems, we lack personal information for 37 persons in the sample (24 from the first PGG session in Faenza and 13 from two PGG sessions, May 25 and 28 in Ragusa). In Model 2 we include a dummy taking value one for these 37 subjects, and zero otherwise (*Missing questionnaire*).

¹⁸As a robustness check, we also run the symmetric regression, where the dependent variable was the frequency of trust (see, Table A-3 in Appendix).

¹⁹We use a Jonckheere-Terpstra test, a non-parametric test for more than two independent samples, designed to test for ordered differences between treatments (Hollander and Wolfe, 1999). For this test, we take as single observation the average contribution at the session level, and the frequency of trustful choices at the matching group level (two in each session).

Table 3: Average individual contributions to the PGG vs. trust at the county level.

| <i>Dep.var.: Average Individual Contribution to the PGG</i> | | |
|---|----------------------|----------------------|
| | Model 1 | Model 2 |
| Trust frequency in the county | 15.320*** (3.421) | 13.310*** (2.817) |
| Male | | 0.309 (0.425) |
| Age 40-59 | | 1.840** (0.791) |
| Age 60 and above | | 0.790 (0.964) |
| Housewife/Retired | | 1.611** (0.559) |
| Student/Unemployed | | 0.492 (0.861) |
| Low understanding | | 0.906 (0.705) |
| Missing questionnaire | | 4.299*** (0.738) |
| N.obs. | 372 | 372 |
| R-squared | 0.027 | 0.087 |

Notes: OLS regression with standard errors robust for clustering at the session level (in parentheses). Symbols ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively.

considered in Table 3. In Model 3 of Table 5, we also add controls for subjects' preferences toward risk of nature: the dummy *Strongly risk averse* (*Risk neutral/Risk loving*) takes value 1 for subjects choosing Lottery 1 or 2 (5 or 6) and 0 otherwise in the Lottery task. In Model 4 of Table 5 we also control for other-regarding preferences – i.e., strong concerns for group wealth and equality. The dummy *Strong concerns for group wealth* is equal to 1 if a subject chooses options 4, 5, or 6 in DG1 and DG2 (Table 2) and zero otherwise. Symmetrically, the dummy *Strong concerns for equality* takes value 1 when a subjects chooses 1, 2 or 3 in DG1 and 4, 5 or 6 in DG3 (Table 2) and 0 otherwise. Results from Tables 4 and 5 indicate that contributions and trust are significantly and positively correlated with latitude.

To further verify that the observed differences in cooperation levels across cities – measured either as contributions to the PGG or as trust in the TG – are not uniquely determined by differences in individual characteristics, in Model 4 of Table 4 and in Models 5 and 6 of Table 5, we remove the coefficient *Latitude* and leave all the aforementioned controls for individual characteristics. Results indicate that age and occupational status may be correlated with trust and contributions, but this effect is orthogonal to the effect of latitude.

Trust levels could however be a response to trustworthiness levels: trustworthiness can influence trust because trust pays off when a high fraction of the population is trustworthy. Being trustworthy also follows a North-South ordering: the individual frequencies of trustworthy choices in the TG are 34% in Cuneo, 33% in Faenza, 32% in Crotona, and 25% in Ragusa (Jonckheere-Terpstra test: p-value=0.049, z=-1.97, n=18, two-sided). However, trustworthiness does not significantly correlate with latitude according to an OLS regression

Table 4: Average individual contribution to the PGG vs. latitude.

| <i>Dep.var.: Average Individual Contribution to the PGG</i> | | | | |
|---|--------------------|---------------------|---------------------|---------------------|
| | Model 1 | Model 2 | Model 3 | Model 4 |
| Latitude | 0.223** (0.077) | 0.182** (0.061) | 0.169** (0.064) | |
| Male | | 0.308 (0.431) | 0.247 (0.446) | 0.234 (0.440) |
| Age 40-59 | | 1.889** (0.802) | 2.004** (0.754) | 2.113** (0.757) |
| Age 60 and above | | 0.836 (0.941) | 0.938 (0.908) | 1.154 (0.881) |
| Housewife/Retired | | 1.614** (0.551) | 1.813*** (0.549) | 1.672** (0.598) |
| Student/Unemployed | | 0.468 (0.863) | 0.665 (0.876) | 0.666 (0.880) |
| Low understanding | | 0.911 (0.720) | 0.969 (0.717) | 0.928 (0.726) |
| Missing questionnaire | | 4.108*** (0.843) | 3.945*** (0.859) | 4.534*** (0.760) |
| Strongly risk averse | | | 0.870 (0.564) | 0.903 (0.572) |
| Risk neutral/Risk loving | | | -0.875* (0.426) | -0.936* (0.432) |
| N.obs. | 372 | 372 | 372 | 372 |
| R-squared | 0.022 | 0.080 | 0.097 | 0.085 |

Notes: OLS regression with standard errors robust for clustering at the session level (in parentheses). Symbols ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively.

(see Table 6). In Model 1, we include each city’s latitude as a regressor, and in Model 2 we add controls for the same individual characteristics considered in Table 3. In Model 3, we also add controls for subjects’ preferences on the risk of nature, and in Model 4 we introduce controls for other-regarding preferences. The average frequency of reciprocal actions is indeed aligned with latitude, but the correlation does not seem to be significant. Differences in trust can indeed be wider than differences in trustworthiness in a population where selfish individuals coexist with altruists and conditional cooperators. Consider that selfish individuals are never trustworthy but could trust. Indeed, trust is profitable in a society with a sufficiently large fraction of altruists and conditional cooperators.

3.3 Can risk preferences explain the cooperation gap?

Trusting in the TG and contributing in the PGG expose participants to the risk of being cheated; hence the cooperation gap can depend on differences in the risk preferences. Unlike cooperation, risk preferences does not exhibit a North-South ordering (see Table 7). Moreover, contributions in PGG and trust in TG strongly correlate with latitude also after controlling for individual risk preferences (see Model 3 in Tables 4 and 5). An Oaxaca decomposition confirms that individual risk preferences does not significantly account for the observed North-South differences in contribution to the PGG and trust in the TG (Tables A-

Table 5: Average individual trust vs. latitude.

| <i>Dep.var.: Trust frequency (at the individual level)</i> | | | | | | |
|--|---------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
| Latitude | 0.018*** (0.005) | 0.017*** (0.005) | 0.018** (0.005) | 0.017** (0.005) | | |
| Male | | -0.006 (0.034) | -0.005 (0.032) | -0.009 (0.030) | 0.003 (0.033) | -0.002 (0.031) |
| Age 40-59 | | -0.075 (0.077) | -0.080 (0.074) | -0.071 (0.072) | -0.081 (0.069) | -0.072 (0.067) |
| Age 60 and above | | 0.033 (0.084) | 0.023 (0.082) | 0.032 (0.081) | 0.040 (0.076) | 0.049 (0.076) |
| Housewife/Retired | | -0.101*** (0.027) | -0.099*** (0.028) | -0.102*** (0.030) | -0.108*** (0.028) | -0.111*** (0.030) |
| Student/Unemployed | | -0.038 (0.061) | -0.038 (0.059) | -0.033 (0.062) | -0.059 (0.059) | -0.053 (0.063) |
| Low understanding | | -0.028 (0.047) | -0.026 (0.047) | -0.030 (0.046) | -0.036 (0.043) | -0.039 (0.044) |
| Missing questionnaire | | -0.046 (0.103) | -0.055 (0.104) | -0.059 (0.106) | -0.139 (0.104) | -0.141 (0.103) |
| Strongly risk averse | | | -0.044 (0.061) | -0.040 (0.057) | -0.035 (0.060) | -0.030 (0.056) |
| Risk neutral/Risk loving | | | -0.005 (0.058) | -0.015 (0.063) | -0.020 (0.059) | -0.030 (0.065) |
| Strong concerns for equity | | | | -0.041 (0.068) | | -0.046 (0.072) |
| Strong concerns for efficiency | | | | 0.017 (0.063) | | 0.019 (0.066) |
| N.obs. | 242 | 242 | 242 | 242 | 242 | 242 |
| R-squared | 0.027 | 0.047 | 0.050 | 0.055 | 0.027 | 0.033 |

Notes: OLS regression with standard errors robust for clustering at the matching-group level (in parentheses). Symbols ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 6: Average individual trustworthiness vs. latitude.

| <i>Dep.var.: Trustworthiness frequency (at the individual level)</i> | | | | | | |
|--|------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
| Latitude | 0.009 (0.008) | 0.006 (0.008) | 0.007 (0.008) | 0.007 (0.008) | | |
| Male | | 0.048 (0.070) | 0.049 (0.070) | 0.049 (0.067) | 0.053 (0.068) | 0.053 (0.066) |
| Age 40-59 | | 0.078 (0.063) | 0.079 (0.062) | 0.079 (0.065) | 0.079 (0.060) | 0.079 (0.064) |
| Age 60 and above | | 0.022 (0.075) | 0.027 (0.079) | 0.027 (0.079) | 0.031 (0.079) | 0.030 (0.079) |
| Housewife/Retired | | 0.042 (0.053) | 0.040 (0.056) | 0.040 (0.057) | 0.038 (0.054) | 0.038 (0.055) |
| Student/Unemployed | | -0.012 (0.067) | -0.008 (0.070) | -0.008 (0.071) | -0.016 (0.066) | -0.016 (0.068) |
| Low understanding | | 0.033 (0.066) | 0.030 (0.072) | 0.030 (0.071) | 0.027 (0.071) | 0.027 (0.071) |
| Missing questionnaire | | 0.005 (0.055) | 0.018 (0.070) | 0.018 (0.068) | -0.014 (0.077) | -0.014 (0.075) |
| Strongly risk averse | | | 0.028 (0.076) | 0.028 (0.078) | 0.031 (0.076) | 0.031 (0.078) |
| Risk neutral/Risk loving | | | 0.037 (0.076) | 0.038 (0.080) | 0.032 (0.075) | 0.032 (0.079) |
| Strong concerns for equity | | | | 0.001 (0.052) | | 0.001 (0.053) |
| Strong concerns for efficiency | | | | -0.002 (0.053) | | -0.000 (0.055) |
| N.obs. | 238 | 238 | 238 | 238 | 238 | 238 |
| R-squared | 0.006 | 0.025 | 0.027 | 0.027 | 0.024 | 0.024 |

Notes: OLS regression with standard errors robust for clustering at the matching-group level (in parentheses). Symbols ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively.

5 and A-6 in Appendix).²⁰ We conclude that the cooperation gap is not driven by differences across cities in aversion to risk.

Table 7: Risk and other-regarding preferences

| | Ragusa | Crotone | Faenza | Cuneo |
|--|--------|---------|--------|-------|
| Tolerance for risk | | | | |
| Strongly risk averse | 28% | 21% | 42% | 25% |
| Weakly risk averse | 44% | 45% | 42% | 47% |
| Not risk averse | 28% | 34% | 16% | 28% |
| | South | ← | → | North |
| Altruism | | | | |
| Strong preference for group wealth | 37% | 37% | 43% | 35% |
| Strong preference for equality | 33% | 37% | 39% | 27% |
| Weak other-regardness or self-interest | 30% | 26% | 18% | 38% |

Notes: We classify 618 participants according to their risk preferences depending on the lottery chosen out of a menu of six (1, 2=not risk averse, 3, 4=weak, 5, 6=strong, Table 1). We also classify 246 participants according to their choices in the three DGs (Table 2): a participant expresses a “strong concern for group wealth” when she favors total surplus in the group both (i) over equality of earnings in the costless choice DG1, and (ii) over self-interest in DG2; by contrast, a participant expresses a “strong concern for equality” when she favors equality of earnings both (i) over total surplus in the costless choice DG1, and (ii) over self-interest in DG3. The third category is residual.

3.4 Can other-regarding preferences explain the cooperation gap?

The observed cooperation gap between North and South of Italy could stem from differences in other-regarding preferences. We detect other-regarding preferences by the individual willingness to pay a cost to increase equality or wealth in the group, and measure it through a set of DGs (see Table 7). The outcome only depends on the dictator; altruists would sacrifice to benefit others, but in DGs there is no scope for strategic interactions or conditional cooperation. The data show that other-regarding preferences cannot explain the cooperation gap. First, unlike cooperation, other-regarding preferences is not aligned along a North-South continuum (see Table 7). Second, individual trust does not significantly correlate with individual levels of other-regarding preferences, and strongly relates with latitude also after controlling for other-regarding preferences and other individual characteristics (see Model 4 in Table 5). Third, an Oaxaca decomposition shows that the North-South difference in trust in the TG is not explained by differences in individual other-regarding preferences (Table A-6 in Appendix). The results at the city level also point in the same direction and suggest that cooperation does not critically depend on other-regarding preferences: for instance, Cuneo displays the weakest concerns for group wealth and yet is the most cooperative county in our sample. We thus conclude that the cooperation gap is for the most part driven factors other than differences in other-regarding preferences across societies.

²⁰ The Oaxaca decomposition (Blinder, 1973; Oaxaca, 1973) is an econometric technique used to decompose a differential between two groups with respect to a variable of interest: in our case, collaboration and trust in Northern and Southern Italy. This technique allows to single out what fraction of this differential can be explained by the available control variable, and what fraction instead remains unexplained.

4 Long-term persistence

Institutions are fundamental in promoting or discouraging cooperation within a society, and they could conceivably be responsible for the persistence of a given cooperative behavior (Zak and Knack, 2001). There are two main mechanisms that could generate such persistence. The first one emphasizes the role of institutions in setting incentives for and payoffs from cooperation, whereby persistence would arise because of institutional path-dependency; current institutions are critical in generating cooperation but they generally change rather slowly over time, hence producing inertia in the observed cooperation levels (Acemoglu and Robinson, 2012). The second mechanism highlights the impact of institutions in shaping individual values and social norms, which have the property of surviving beyond the life of the parent institution, thus sustaining the persistence of a given cooperative behavior (Putnam et al., 1993; Guiso et al., 2004; Tabellini, 2008; Guiso et al., 2008).

Institutional path-dependency cannot account for our findings: as a result of the experimental methodology, participants across Italy faced no differences across locations in terms of institutions and of payoffs from cooperation, thus avoiding elements that could potentially shape behavior such as the degree of corruption, the efficiency of the public administration or the level of mafia intrusion. The fact that 150 years of common history in terms of formal institutions and language have failed to achieve full convergence in cooperation levels suggests that the source of persistence is to be sought in values and norms.

Here we focus on the second mechanism and study whether dissimilar pre-unification historical experiences, up to a thousand years ago, could have shaped local values and norms, which can in turn account for differences in current behavior. There is only so much that we can test with four cities as to what could have brought about different cooperative dispositions; still, variability within this small set is such as to make it worthwhile to explore its origins. We consider both political institutions and wars.

Table 8: Historical roots of cooperation

| | Ragusa | Crotone | Faenza | Cuneo |
|---|--------|---------|--------|-------|
| | South | ← | → | North |
| Quality of political institutions (average score) | 1.86 | 1.59 | 2.55 | 2.43 |
| Number of external violent conflicts | 1 | 5 | 8 | 12 |

Notes: Data refer to years 1000-1860 for institutions and 1000-1849 for conflicts, and starts in 1175 for Cuneo (foundation year). The quality of institutions reflects the POLITY IV dataset definition of “constraints to the executive” and ranges from 1 (=unlimited authority) to 7 (=accountable executive, constrained by checks and balances). External violent conflicts are defined as the sum of offensive warfare, defensive warfare, and army stationing.

Robert Putnam argued that the free city-state experience in Italy, between 1000 and 1300 A.D., could be the source of the North-South gap in civic norms (Putnam et al., 1993; Guiso et al., 2004). Other scholars have extended the conjecture to the quality of political

institutions more generally as the force shaping people’s values and norms (Tabellini, 2010; Guiso et al., 2008). For this reason, we focus on the institutions protecting the rule of law in each city between 1000 and 1850. We followed the approach of Tabellini (2008) in codifying the constraints on the executive power and extended his evidence by six centuries (Table A-7 and A-2 in Appendix).²¹ The score we obtained in the quality of institutions displays the same North-South gap as our experimental measures of the ability to cooperate, but does not significantly differentiate between the two pairs of cities (see, Table 8).

The attempt to explain cooperative dispositions with institutional differences inevitably begs further questions, as it makes one wonder where do these institutions come from. A number of theories have taken the matter to a more fundamental level and claimed that human groups and their institutions are differentially shaped as a result of varying exposure to violent external conflicts (Tilly, 1992; Bowles, 2009). Gene-culture co-evolutionary models posit that external conflicts select groups with a higher frequency of cooperators, since groups so configured have a higher probability of prevailing over an enemy (Bowles, 2009). As a result, societies with a history of external conflicts would be characterized by a higher level of in-group cooperation. Pursuing this conjecture, we coded violent external conflicts in each city since the year 1000 (for further details, see Appendix 3.2). The higher frequency of conflicts in the North compared to the South matches not only the gap in cooperation levels in the two broad regions, but also the ranking of cooperation levels that we found in our four cities (Tables A-8 and A-9). We can thus surmise that conflicts which occurred during the eight centuries before Unification in 1861 have an impact on the present cooperative norms of societies. Should this correlation survive a stricter test based on a larger sample of cities, this evidence would show the genesis of group differences from long-run collective experiences.

5 Discussion and Conclusions

Here we uncover the behavioral foundations of a classic disparity in economic development, civic norms, and quality of political institutions between the North and the South of Italy. In a field experiment with a representative sample of the general population, we find that the ability to cooperate is stronger in the North than in the South. This cooperation gap emerges even though our methodology silences the impact of legal and economic institutions; moreover, everyone faces identical payoffs from cooperation, thus these factors cannot be blamed. Cooperation is independently measured as “trust” and as “contributions to a common project” and both measures yield the same ranking in the ability to cooperate. The empirical alignment between trust and contributions level is novel and it also displays the robustness of the behavioral difference in cooperation between the North and the South of Italy.

²¹ For a detailed discussion of the coding procedure, see Appendix 3.1

The ability to cooperate turned out to be unrelated to individual characteristics, such as other-regarding or risk preferences: people in the South are no less altruistic or risk tolerant than people in the North. These results thus suggest that the gap is not due to “moral flaws”, as famously claimed by Banfield and Fasano (1958). Although in principle other-regarding and risk preferences could be a determinant of cooperation, in this case study their empirical impact is dwarfed by other factors. Our interpretation is that the North-South gap originates in different abilities to cooperate, which represent emergent properties of a society. Two societies with similar risk preferences and similar fractions of altruists could achieve different in-group cooperation levels. One possible explanation for the North-South gap is given by the differences in social norms of conditional cooperation and in the expectations thereof. Cooperation is unconditional when it is supported by altruists who do not act for a personal gain; however, it can also emerge in a population of conditional cooperators, who respond to the expectation – or the realization – that others are cooperative.

This study carries policy implications. If the North-South gap in the ability to cooperate is rooted in social norms, institutional change alone is unlikely to close the gap. One would first need to understand the origin of such distinct norms at the local level. Regional disparities in Italy survived 150 years of common national history, thus displaying long-term persistence, and several scholars have pointed at heterogeneous experiences in the distant past as the source of the gap in cooperation levels Putnam et al. (1993); Nunn and Wantchekon (2011); Guiso et al. (2008). We collected a millennium-long time span dataset, which suggests that the most plausible conjecture in terms of historical experiences is related to the frequency of violent conflicts against external enemies Bowles (2009); Tilly (1992). The need to react to external threats may positively select norms of conditional cooperation in a society, and lay the bases for a more cooperative future.

Acknowledgments: We thank A. Barr, S. Bowles, C.F. Camerer, T.N. Cason, R. Durante, B. Herrmann, and D.G. Rand for critical discussion and reading of the manuscript, S. Amadori for historical research, and M. Adami for technical support and help in conducting the experiments. We also thank P. Mantovani, P. Motetti, G. Pignataro, C. Tagliapietra for logistic help. M. Casari gratefully acknowledges support from an ERC Starting Grant (No. 241196).

References

- Acemoglu, D. and J. Robinson (2012). *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. Crown Publishing Group.
- Banfield, E. C. and L. Fasano (1958). *The Moral Basis of a Backward Society*. The Free Press.
- Benhabib, J., M. O. Jackson, and A. Bisin (2011). *Handbook of Social Economics*. Handbooks in Economics, New York, Elsevier.
- Bigoni, M., S. Bortolotti, M. Casari, and D. Gambetta (2013). It takes two to cheat: An experiment on derived trust. Mimeo.
- Blinder, A. S. (1973). Wage Discrimination: Reduced Form and Structural Estimates. *The Journal of Human Resources* 8(4), 436–455.
- Bolton, G. E. and A. Ockenfels (2000). ERC: A Theory of Equity, Reciprocity, and Competition. *American Economic Review* 90(1), 166–193.
- Bowles, S. (2009). Did Warfare Among Ancestral Hunter-Gatherers Affect the Evolution of Human Social Behaviors? *Science* 324(5932), 1293–1298.
- Bowles, S. and H. Gintis (2011). *A Cooperative Species: Human Reciprocity and Its Evolution*. Princeton University Press.
- Buchan, N., G. Grimalda, R. Wilson, E. Brewer, M. and Fatas, and M. Foddy (2009). Globalization and Human Cooperation. *Proceedings of the National Academy of Sciences* 106(11), 4138–4142.
- Cartocci, R. (2007). *Mappe del tesoro: atlante del capitale sociale in Italia*. Il mulino.
- Cole, J. and E. Wolf (1974). *The Hidden Frontier: Ecology and Ethnicity in an Alpine Valley*. University of CALIFORNIA Press.
- Daniele, V. and P. Malanima (2011). *Il divario Nord-Sud in Italia 1861-2011*. Rubbettino.
- Eckel, C. and P. Grossman (2008). Forecasting risk attitudes: An experimental study using actual and forecast gamble choices. *Journal of Economic Behavior & Organization* 68(1), 1 – 17.
- Fehr, E. and K. Schmidt (1999). A Theory Of Fairness, Competition, And Cooperation. *The Quarterly Journal of Economics* 114(3), 817–868.
- Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10(2), 171–178.
- Galasso, G. (1980). *Storia d'Italia*. UTET.

- Gneezy, A. and D. Fessler (2012). Conflict, sticks and carrots: war increases prosocial punishments and rewards. *Proceedings of the Royal Society B: Biological Sciences* 279(1727), 219–223.
- Guiso, L., P. Sapienza, and L. Zingales (2004). The Role of Social Capital in Financial Development. *The American Economic Review* 94(3), pp. 526–556.
- Guiso, L., P. Sapienza, and L. Zingales (2006). Does Culture Affect Economic Outcomes? *Journal of Economic Perspectives* 20(2), 23–48.
- Guiso, L., P. Sapienza, and L. Zingales (2008). Alfred Marshall Lecture Social Capital as Good Culture. *Journal of the European Economic Association* 6(2-3), 295–320.
- Guiso, L., P. Sapienza, and L. Zingales (2010). Civic Capital as the Missing Link. EIEF Working Papers Series 1005, Einaudi Institute for Economic and Finance (EIEF).
- Henrich, J., J. Ensminger, R. McElreath, A. Barr, C. Barrett, A. Bolyanatz, J. Cardenas, M. Gurven, E. Gwako, N. Henrich, C. Lesorogol, F. Marlowe, D. Tracer, and J. Ziker (2010). Markets, Religion, Community Size, and the Evolution of Fairness and Punishment. *Science* 327(5972), 1480–1484.
- Herrmann, B., C. Thöni, and S. Gächter (2008). Antisocial punishment across societies. *Science* 319(5868), 1362.
- Hollander, M. and D. Wolfe (1999). *Nonparametric Statistical Methods, 2nd Edition* (2 ed.). Wiley-Interscience.
- Knack, S. and P. Keefer (1997). Does Social Capital Have an Economic Payoff? A Cross-Country Investigation. *The Quarterly Journal of Economics* 112(4), 1251–1288.
- La Porta, R., F. Lopez-de-Silanes, A. Shleifer, and R. Vishny (1997). Trust in Large Organizations. *American Economic Review* 87(2), 333–38.
- Nowak, M. and R. Highfield (2011). *SuperCooperators: Altruism, Evolution, and Why We Need Each Other to Succeed*. Free Press.
- Nunn, N. and L. Wantchekon (2011). The Slave Trade and the Origins of Mistrust in Africa. *American Economic Review* 101(7), 3221–52.
- Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review* 14(3), 693–709.
- Putnam, R., R. Leonardi, and R. Nanetti (1993). *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton University Press.
- Tabellini, G. (2008). “Institutions and Culture” Presidential lecture. *Journal of the European Economic Association* 6, 255–294.

- Tabellini, G. (2010). Culture and Institutions: Economic Development in the Regions of Europe. *Journal of the European Economic Association* 8(4), 677–716.
- Tilly, C. (1992). *Coercion, capital, and European states, AD 990-1992*. Studies in social discontinuity. Blackwell.
- Voors, M., E. Nillesen, P. Verwimp, E. Bulte, R. Lensink, and D. V. Soest (2012). Violent Conflict and Behavior: A Field Experiment in Burundi. *American Economic Review* 102(2), 941–64.
- Zak, P. and S. Knack (2001). Trust and Growth. *Economic Journal* 111(470), 295–321.

Appendixes

1 Selection of locations and participants

We selected four medium-sized cities in Italy locations to conduct our experiments. Table A-1 summarizes the main characteristics of the four cities and Figure A-1 depicts the laboratory in the different locations.

Study participants were recruited among the general population of each county. Table A-2 summarizes the socio-demographic characteristics of the actual sample.

Two professional companies -Metis-Ricerche and Demoskopea- were hired for the recruitment of subjects; we provided them with a script to approach potential participants.¹ The recruiters had no prior knowledge of the purpose or content of the study. We asked them to recruit people residing in or outside the town where we ran the experiment. In addition to the aforementioned requirements, special categories of people were ex-ante barred from participation: employees of the research sector; people who took part in market researches in the preceding three months prior to the contract; family members of the recruiters; employees of marketing companies and of the press in general. Moreover, no more than two people per session should be acquaintances each other.

One company (Metis-Ricerche) recruited subjects for the first five sessions in Faenza. Potential subjects were identified by using telephone book entries and they were approached by telephone calls. All phases of the recruitment process were carried out from the headquarters of the company, and, in case of acceptance, the company mailed a confirmation letter to the participant. Metis-Ricerche decided not to renew the contract for the other locations because the recruitment procedures turned out to be more costly than expected for them. The other company (Demoskopea) recruited subjects for two sessions in Faenza and for five sessions in each of the other cities. Local representatives of Demoskopea carried out the recruitment of the subjects in each county. Local representatives of Demoskopea carried out the recruitment of the subjects in each county, by choosing/ randomly extracting names from telephone books and by using random contacts obtained through personal interactions, as instructed by the headquarter.

¹The script is available upon request to the authors.

Table A-1: Characteristic of the selected counties

| City | South | | | North | | | Source | Year |
|--|---------|---------|---------|---------|-------------------------------|-----------|--------|------|
| | Ragusa | Crotone | Faenza | Cuneo | | | | |
| Population | | | | | | | | |
| City population | 73,333 | 61,392 | 57,664 | 55,464 | Istat | 2009 | | |
| County population | 316,113 | 173,812 | 389,509 | 589,586 | Istat | 2009 | | |
| Population density (residents/squared km) | 196 | 101 | 209 | 85 | Istat | 2009 | | |
| Foreign citizens (% of residents) | 2.8% | 1.5% | 4.5% | 4.6% | Istat | 2003 | | |
| Geography | | | | | | | | |
| Latitude | 36°56'N | 39°05'N | 44°17'N | 44°23'N | | | | |
| City elevation (mt on the sea level) | 520 | 8 | 34 | 534 | Sopros | | | |
| Economy | | | | | | | | |
| Per-capita GDP (euros) | 20,008 | 15,447 | 29,110 | 30,414 | Unioncamere | 2008 | | |
| Regional cost of living (Italy=100) | 90.9 | 83.9 | 105.2 | 103.1 | | | | |
| Active labor force (% of population above 15) | 43.8% | 32.6% | 53.8% | 54.9% | Istat | | | |
| Social capital | | | | | | | | |
| Number of voluntary associations/1,000 inhabitants | 37.77 | 16.08 | 64.45 | 33.73 | Istat | 2000 | | |
| (<i>ranking out of 103</i>) | (44) | (101) | (9) | (48) | | | | |
| Referenda turnout as % of eligible to vote | 78% | 64% | 91% | 83% | Data base Guiso et al. (2010) | 2004 | | |
| (<i>ranking out of 103</i>) | (66) | (101) | (2) | (56) | | | | |
| N. of blood donations per 1000 inhabitants | 64.5 | 18 | 84.8 | 34.2 | Cartocci (2007) | 2002 | | |
| (<i>ranking out of 103</i>) | (5) | (95) | (1) | (58) | | | | |
| Justice | | | | | | | | |
| Average duration of a civil trial | 1007 | 1463 | 1004 | 555 | Ministry of Justice | 2006 | | |
| Violent deaths (N. homicides/100,000 inhabitants) | 1.93 | 5.8 | 0.53 | 0.17 | Istat | 2007 | | |
| Education | | | | | | | | |
| 6th grade standardized test in Reading and Mathematics | 99 | 101 | 28 | 16 | Invalsi | 2009/2010 | | |
| (<i>rankings out of 103</i>) | (94) | (100) | (25) | (6) | | | | |
| Literacy in the region | 97.20% | 95.30% | 99.30% | 99.30% | Istat | 2001 | | |

Table A-2: Socio-demographic variables

| | Ragusa | Crotone | Faenza | Cuneo | Total |
|---------------------------------|---|--|---|--|-------|
| Male | 43.9% | 47.9% | 49.4% | 49.3% | 47.7% |
| Age | | | | | |
| 18-39 | 36.7% | 23.2% | 30.1% | 23.6% | 28.4% |
| 40-59 | 40.3% | 53.5% | 42.3% | 45.1% | 45.3% |
| 60 or + | 23.0% | 23.2% | 27.6% | 31.3% | 26.3% |
| Employment Status | | | | | |
| Employed | 34.8% | 42.3% | 45.5% | 42.4% | 41.4% |
| Housewife/Retired | 34.8% | 30.3% | 27.6% | 38.9% | 32.8% |
| Student/Unemployed | 30.4% | 27.5% | 26.9% | 18.8% | 25.9% |
| Education Level | | | | | |
| 8th grade or lower | 43.8% | 43.7% | 23.1% | 24.3% | 33.3% |
| High school | 40.9% | 45.8% | 45.5% | 45.8% | 44.6% |
| College/Master/Ph.D. | 15.3% | 10.6% | 31.4% | 29.9% | 22.1% |
| Rootedness | | | | | |
| Elementary school in the county | 99.3% | 94.4% | 91.7% | 99.3% | 96.0% |
| Mother born in the county | 90.7% | 87.3% | 75.6% | 85.4% | 84.5% |
| Father born in the county | 90.6% | 85.2% | 71.2% | 72.2% | 79.5% |
| Sessions | | | | | |
| PGG sessions (year 2011) | 26/05 6.30 pm 27/05 6.30 pm 28/05 4.00 pm | 19/05 6.30 pm 20/05 6.30 pm 21/05 10.30 am | 02/03 6.30 pm 04/03 6.30 pm 05/03 3.30 pm 01/10 10.30 am | 10/06 6.30 pm 11/06 10.30 am 11/06 3.30 pm | – |
| TG sessions (year 2011) | 25/05 6.30 pm 28/05 10.30 am | 21/05 3.30 pm 21/05 7.00 pm [†] | 03/03 6.30 pm 05/03 10.30 am 01/10 3.30 pm | 08/06 6.30 pm 09/06 6.30 pm | – |
| Subjects | | | | | |
| N. of participants | 152 | 142 | 180 | 144 | 618 |
| N. of completed questionnaires | 139 | 142 | 156 | 144 | 581 |

[†]For technical problems, one session (i.e., Crotone, 18/05/2011) was stopped after task two and the session was rescheduled; 15 subjects from the original session showed-up for completing the experiment on a later date (i.e., Crotone 21/05/2011, 7pm) and 12 additional subjects were recruited; only task 3 and 4 were played in that session.

Figure A-1: Mobile laboratory

(a) Crotone



(b) Cuneo



(c) Faenza



2 Additional statistical analysis

This section presents more statistical analysis and details on the relation between contributions to the PGG, trust, trustworthiness, other-regardness, risk of nature, and demographic characteristics of the subject pool.

In Table 3 in the main text, we provide evidence of a positive and significant relation between individual contributions and trust. As a robustness check, we also run the symmetric linear regression, where the dependent variable is the frequency of trust measured at the individual level (Table A-3), and the only independent variable in Model 1 is the average contribution to the public good at the county level. Results indicate that an increase by 1 point (5 percent) in the average contribution to the PGG corresponds to a highly significant increase of the frequency of trust, by about 6 percentage points. In a second specification (Table A-3, Model 2), we include additional regressors to control for individual characteristics. Results from Model 2 confirm that the positive and significant relation between trust and individual contributions to the public good is not driven uniquely by differences in the individual characteristics of our subjects across counties.

Table A-3: Average individual trust vs. contribution to the PGG at the county level.

| <i>Dep.var.: Trust frequency (at the individual level)</i> | | |
|--|---------------------|----------------------|
| | Model 1 | Model 2 |
| Average Contribution to the PGG in the county | 0.060*** (0.011) | 0.065*** (0.015) |
| Male | | -0.008 (0.033) |
| Age 40-59 | | -0.079 (0.073) |
| Age 60 and above | | 0.044 (0.083) |
| Housewife/Retired | | -0.110*** (0.026) |
| Student/Unemployed | | -0.041 (0.060) |
| Low understanding | | -0.034 (0.046) |
| Missing questionnaire | | -0.029 (0.108) |
| N.obs. | 242 | 242 |
| R-squared | 0.023 | 0.047 |

Notes: OLS regression with standard errors robust for clustering at the matching-group level (in parentheses). Symbols ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively.

In Table A-4 we test whether a significant relation emerges between latitude and the average gains from trust, measured as the average profit obtained by a subject in all periods when he had the role of truster and decided to trust. Latitude is not significantly correlated with the gains from trust.

Table A-4: Average individual gains from trust vs. latitude.

| <i>Dep.var.: Average Gains from Trust</i> | |
|---|------------------|
| Latitude | 0.391 (0.260) |
| N.obs. | 176 |
| R-squared | 0.007 |

Notes: OLS regression with standard errors robust for clustering at the matching-group level (in parentheses). Symbols ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively.

As a further test, we measure the role of individual factors such as the difference in risk and other-regarding preferences in explaining the differential in collaboration and trust in the North and South of Italy by means of a (Blinder-) Oaxaca decomposition (Blinder, 1973; Oaxaca, 1973). The Oaxaca decomposition is an econometric technique used to decompose a differential between two groups with respect to a variable of interest: in our case, collaboration and trust in Northern and Southern Italy. This technique allows to single out what fraction of this differential can be explained by the available control variable, and what fraction instead remains unexplained. This technique can be implemented only with respect to two groups. Hence, we can only decompose the North-South differential in cooperation, but not the effect of latitude across the four cities.

The results of the Oaxaca decomposition for the Public Goods experiment are in Table A-5 and for the Trust Game experiment are in Table A-6. In the PGG, the estimated difference in the average contribution between Northern and Southern counties is equal to -1.213 points (Prediction South - Prediction North), and it is statistically significant at the 5% level (Table A-5, Model 1). One can see how much of this difference can be attributed to differences in risk preferences. This dimension is coded through two dummy variables (*Strongly risk averse and Risk neutral/Risk loving*) and can explain -0.111 point difference in average contribution, which amounts to only 9%.

Similarly, in the TG, the estimated difference in average trusting behavior is equal to -0.111 and is statistically significant at the 1% level (A-6, Model 2). For the TG, we have both individual measures of risk and other-regarding preferences (i.e., the dummies Preference for Equity and Preference for Efficiency). These two dimensions taken together cannot explain any of the point difference in average trusting behavior. Indeed, the sign of the coefficients points in opposite directions (0.002).

Table A-5: Oaxaca decomposition of North South difference in average contribution to the public good.

| <i>Dependent variable: contribution to the PGG</i> | | |
|--|---------------------|---------------------|
| | Model 1 | Model 2 |
| Prediction South | 8.153*** (0.378) | 8.153*** (0.358) |
| Prediction North | 9.366*** (0.440) | 9.366*** (0.425) |
| Difference | -1.213** (0.580) | -1.213** (0.556) |
| Strongly risk averse | -0.067 (0.062) | -0.076 (0.064) |
| Risk neutral/Risk loving | -0.044 (0.043) | -0.059 (0.054) |
| Explained part | -0.111 (0.081) | 0.024 (0.149) |
| Male | | -0.011 (0.033) |
| Low understanding | | 0.006 (0.043) |
| Age 40-59 | | 0.058 (0.100) |
| Age 60 or above | | 0.002 (0.017) |
| Housewife/Retired | | 0.104 (0.094) |
| Unexplained part | -1.102* (0.566) | -1.237** (0.518) |

Notes: Regression with standard errors robust for clustering at the session level (in parentheses). Symbols ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively.

Table A-6: Oaxaca decomposition of North South difference in average trusting behavior.

| <i>Dependent Variable: contribution to the TG</i> | | | |
|---|----------------------|----------------------|----------------------|
| | Model 1 | Model 2 | Model 3 |
| Prediction South | 0.364*** (0.009) | 0.364*** (0.009) | 0.364*** (0.013) |
| Prediction North | 0.475*** (0.028) | 0.475*** (0.029) | 0.475*** (0.033) |
| Difference | -0.111*** (0.029) | -0.111*** (0.030) | -0.111*** (0.035) |
| Strongly risk averse | 0.007 (0.010) | 0.006 (0.009) | 0.006 (0.009) |
| Risk neutral/Risk loving | -0.001 (0.009) | -0.003 (0.010) | -0.001 (0.010) |
| Explained part | 0.006 (0.010) | 0.002 (0.010) | 0.003 (0.015) |
| Strong concerns for equality | | -0.001 (0.003) | -0.001 (0.003) |
| Strong concerns for efficiency | | -0.000 (0.002) | -0.000 (0.002) |
| Male | | | 0.000 (0.004) |
| Low understanding | | | -0.000 (0.002) |
| Age 40-59 | | | -0.002 (0.005) |
| Age 60 or above | | | -0.008 (0.012) |
| Housewife/Retired | | | 0.009 (0.008) |
| Unexplained part | -0.117*** (0.031) | -0.113*** (0.031) | -0.115*** (0.036) |

Notes: Regression with standard errors robust for clustering at the session level (in parentheses). Symbols ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively.

3 Historical roots of cooperation

A growing number of scholars from different disciplines have interpreted cooperation, and more generally economic development, as a by-product of distant historical events. In the following, we link present cooperation levels as measured in our field experiment to data about the quality of past institutions and the intensity of violent external conflicts in the last thousand years.

3.1 Institutions and cooperation

Political scientists have pointed out that institutions can shape culture. In particular, the ongoing presence of well functioning institutions may promoted cultural traits, such as trust and cooperation, commonly viewed as favorable to economic growth and development. More specifically, enforcing the rule of law may have prevented abuses against personal freedoms and property rights, and it might have protected cooperators from the risk of exploitation. Table A-7 reports data on the level of constraints on the executive authority in our four locations from the year 1000. Better institutions are more accountable, and they are constrained by checks and balances that limit the arbitrary use of force and privileges. Republican institutions – where the rule of law and property rights are respected – are understood to foster generalized morality and trust. The quality of institutions ranges from 1 (unlimited authority) to 7 (accountable executive, constrained by checks and balances).²

We extended Tabellini (2010)’s codification back in time, to include years from 1000 to 1580. For the years before 1580, we considered one century as the reference time span for coding institutions, as detailed information on shorter time spans were harder to evaluate. We followed as close as possible the criteria set in POLITY IV for the variable *xconst*³ and we also adapted by Tabellini when coding the quality of institutions from 1580-1860. As data source we relied on Galasso (1980) as data source. The political entity taken as a reference is not the city itself, but the State comprising the city.cities⁴

Table A-7 summarizes the scores from our own research, Tabellini, and POLITY IV, covering a total of 1012 years. The aggregate score for a city is computed as the average of each score weighted by its respective time span in years. Table A-7 reports two aggregate scores, one for the years up to the Italian Unification (1861) and one for the whole period (years 1000-2012). The data show a gap in the quality of past political institutions between the North and the South. The city ranking sees Faenza at the top, while Cuneo, Ragusa, and Crotone at the bottom.

The quality of institutions in the distant past may have had a lower impact on current values and norms than that of institutions in the recent past. To take this into account, we applied

²Further details about the coding of political institutions and data sources can be found in Tabellini (2010).

³<http://www.systemicpeace.org/polity/polity4.htm>

⁴An experienced historian, Saverio Amadori, unaware of the goal of the study, was hired to carry out the search and classification of the quality of institutions. Only edited sources were used.

a discounting factor to calculate an aggregate score for each city. In particular, we divided the time span into 25-year intervals, which roughly correspond to a generation. Consider a discount factor of 0.95 from one generation to the next, and so on; this level of discounting gives, for instance, a weight of 1 to the quality of institutions in 1999-1975 and a weight of $(0.95)^{(500/25)} = 0.36$ to the quality of institutions in years 1499-1475. As a robustness check, Figure A-2 illustrates the aggregate scores by city for different levels of discounting, ranging from 1 through 0.7. Panel *a* of Figure A-2 includes data up to 1849 and Panel *b* up to 1999. City ranking is generally robust to different levels of discounting of past events, with the notable exception of Cuneo that overtakes Faenza for lower discount factors.

Table A-7: Quality of institutions in four Italian cities: Years 1000-2012.

| Source | Years | Time span (years) | Ragusa | Crotone | Faenza | Cuneo |
|-----------------------------|--------------|----------------------|-------------------|---------|--------|-------|
| | | | Score | | | |
| Own research | XI century | 100 | 2 | 2 | 2.5 | |
| Own research | XII century | 100 | 2.5 | 2.5 | 4 | 3 |
| Own research | XIII century | 100 | 3 | 2 | 2.5 | 4 |
| Own research | XIV century | 100 | 2 | 2 | 3 | 2 |
| Own research | XV century | 100 | 2 | 1.5 | 3 | 3 |
| Own research | XVI century | 80 | 2 | 1 | 2 | 2.5 |
| Tabellini | 1580-1620 | 70 | 1 | 1 | 2 | 2 |
| Tabellini | 1680-1720 | 74 | 1 | 1 | 2 | 2 |
| Tabellini | 1730-1770 | 50 | 1 | 1 | 2 | 1 |
| Tabellini | 1780-1815 | 41 | 1 | 1 | 2 | 1 |
| POLITY IV | 1816-1847 | 32 | 1 | 1 | 1 | 1 |
| POLITY IV | 1848-49 | 2 | 1 | 1 | | 3 |
| POLITY IV | 1849-60 | 11 | 1 | 1 | 1 | 3 |
| POLITY IV | 1861-1900 | 40 | 3 | 3 | 3 | 3 |
| POLITY IV | 1901-21 | 21 | 5 | 5 | 5 | 5 |
| POLITY IV | 1922-27 | 6 | | | | |
| POLITY IV | 1928-42 | 15 | 1 | 1 | 1 | 1 |
| POLITY IV | 1943-48 | 6 | | | | |
| POLITY IV | 1949-2012 | 64 | 7 | 7 | 7 | 7 |
| | | | Weighted averages | | | |
| Average for years 1000-1860 | | | 1.86 | 1.59 | 2.55 | 2.43 |
| Average for years 1000-2012 | | | 2.21 | 1.98 | 2.80 | 2.74 |

Notes: Empty spaces denote either non existence for the city (Cuneo in XI century) or data not available in the POLITY IV dataset. Cuneo was founded in 1175 and hence the weighted averages comprises a smaller number of years: 685 for preunification history and 837 for whole period.

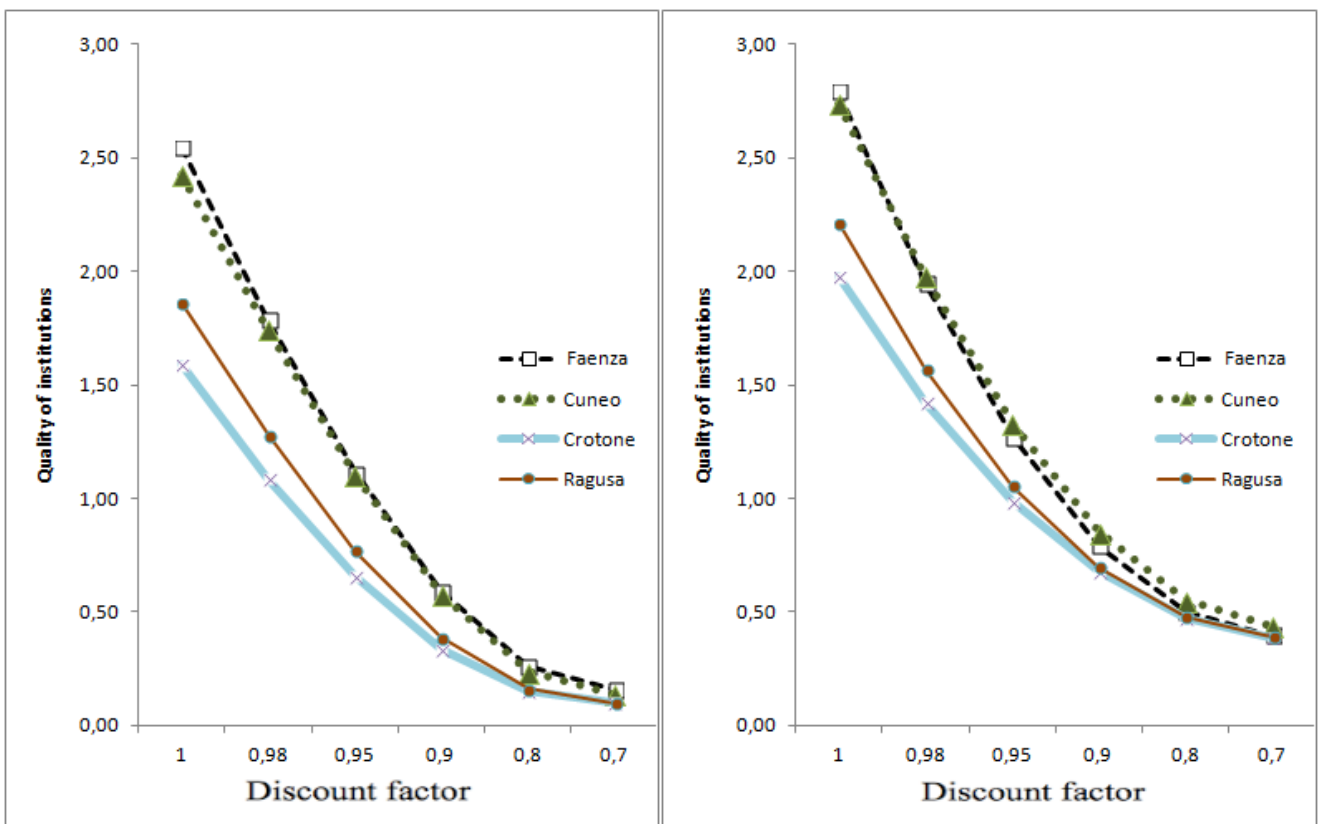
3.2 External violent conflicts and cooperation

A growing number of theories have acknowledged the role of violent conflicts in shaping cooperative attitudes. For instance, Tilly (1992) conjectures that a history of violent conflicts has a key impact on the internal social and political dynamics of societies. This author studies the development of the modern national states in Europe between the year 990 and 1992, and he argues that political rules and apparatuses evolved, together with social and civic norms, as a by-product of the pursuit of war and military capacity. Bowles and Gintis (2011) posit that conflicts generate a need for cooperation within the group because the higher the internal cohesion, the higher the probability of prevailing over the enemy. Winners, in turn, take over

Figure A-2: Quality of political institutions for different levels of discounting

(a) Years 1000-1849

(b) Years 1000-1999



the power and the resources of the losers, hence granting higher chances of survival to their offspring, and spreading their culture and institutions. As a result, the emerging society will be characterized by a higher level of cooperation. Indeed, archaeological and ethnographic evidence on warfare from Late Pleistocene and early Holocene suggests that selective pressure of conflicts was strong enough to select and sustain cooperative norms within human societies (Bowles, 2009). There is also evidence that pro-social behavior within the same generation is enhanced by exposure to warfare. Cooperation levels in a laboratory experiment were indeed higher during the 2006 Israeli-Hezbollah war, compared to before or after the war itself (Gneezy and Fessler, 2012). In another experiment led in Burundi, subjects who reported to be exposed to violence during the 1993-2003 Hutu-Tutsi war were more likely to behave altruistically toward their neighbors in an experiment (Voors et al., 2012).

In line with these theories, we conjecture that a higher frequency of past conflicts in a particular area positively affects present in-group cooperative behavior among the people who live in that area. Our analysis differs from the existing literature because it focuses on the link between *current* levels of cooperation – measured by means of experiments – and the historical measures of violent external conflicts’ incidence at the local level.

3.3 Coding violent conflicts in Italy

We coded external violent conflicts in the four cities⁵ over a time span of more than eight centuries: from 1000 to 1861 (Italian unification).⁶ This time span is similar to the one analyzed by Tilly: however, we do not consider the years following the Italian unification, because from that year on our four cities experienced common history and institutions. We considered the number of conflicts at the city level. Obtaining a county-level measure of conflicts was not feasible, since the boundaries of the counties changed repeatedly during the considered time interval (Table A-10). We examined violent events reported in edited volumes, and considered as conflicts only events above a given intensity threshold. The intensity of an event was mainly determined by the war effort exerted by the city and by the impact of the war on the city. More specifically, the following criteria were taken into account to evaluate the intensity of each causality: duration of the event; its impact given the number of soldiers; number of casualties and injuries; severity of the damage to civilian and military structures inside and outside the city walls; bombing; city looting; devastation of the countryside in terms of field harvests and livestock; amount of monetary and territorial compensations; and speed of the attack. An experienced historian evaluated each event based on the aforementioned criteria. We considered three types of violent conflicts: defensive warfare, offensive warfare, army stationing.

⁵An experienced historian, Saverio Amadori, unaware of the goal of the study, was hired to carry out the search and classification of conflicts. Only edited sources were used.

⁶Faenza, Crotona, and Ragusa were founded before 1000. Cuneo instead was founded 1175 and hence previous conflicts were not coded for this city.

Defensive warfare is a violent conflict sustained by the city (and thus promoted by the enemy) such as a siege, an invasion of the countryside, or occupation of buildings or fortifications in the territory under the political rule of the city. Such an event is classified as defensive warfare when there is city looting, or in presence of casualties or injuries. In alternative, an event is classified as defensive warfare when two or more of the following elements are simultaneously present: (a) destruction of the territories surrounding the city or stationing of armies; (b) damage to civilian and military facilities; (c) bombing; and (d) surprise attack (without a prior declaration of war). This category can include conflicts sustained by the political entity in which the city is located (e.g., the Kingdom of Naples for Crotone), but only if the conflict somehow involved the city itself. It also includes warfare episodes fought by opposing factions inside the city itself, but only when there is an intervention by supporting external military forces coming from other cities. On the other hand, our definition of violent conflicts excludes purely internal disputes which involve different factions of the same city, conspiracies, hostages handovers, coups d'état and internal changes in the control of political power, or disputes that were resolved simply by means of monetary payments. We also exclude the transit of armies if the army abstains from damages, even in exchange for monetary or territorial compensations.

Offensive warfare is a violent conflict with casualties and injuries promoted by the city, such as the siege of another city, the invasion of the countryside, occupation of buildings or fortifications in the territory of other entities. Such an event is classified as offensive warfare where there are (a) casualties or injuries among the city population excluding the soldiers of fortune and, in addition, there is either an (b) army stationing, or (c) bombing. Financial and logistic support to warfare fought by mercenaries and soldiers of fortune can be part of the offensive warfare. Dispatch of troops by the city's political power, even when is alliance with other cities or states, is also considered conflict. We exclude conflicts that only entail taking hostages among the enemies, and also conflicts promoted by the political entity in which the city is located, if the city itself is neither the decision-making seat nor has the logistic management of the conflict, and if the conflict does not include military service for its own citizens. We also exclude disputes resolved simply by monetary transactions or territorial compensations, and those that involved only army stationing.

Army stationing is the presence of others' armies in the area under the political rule of the city, even in the absence of an attack. The criteria for identification of an event as a conflict comes from the combination of three elements: (a) the size of the foreign army; (b) the length of the stationing; (c) and the explicit will to damage the territory. We exclude the simple transit of an army. If the transit includes raids, then it is classified as defensive warfare.

3.3.1 Ranking cities by level of violent conflicts

Based on the coding described above, we built several measures of conflict intensity (Tables A-8 and A-9). The first measure is the raw sum of all conflicts classified as defensive or offensive warfare, and army stationing. This measure of conflict intensity displays a latitudinal ordering; Cuneo is the city with the highest conflict intensity, followed by Faenza, Crotone, and Ragusa. It is worth noticing that Cuneo has 12 times as many conflicts as Ragusa, and more than twice as Crotone. The difference is even more striking if we consider that observations for Cuneo are available starting only from 1175. The ranking does not change when we consider the average number of conflicts per generation – conventionally taken as covering 25 years – which is equal to the total number of conflicts divided by the number of generations for which data are available.

Conflicts are not evenly distributed over time in the four cities and it is possible that more recent conflicts might have had a higher impact on present social norms and behavior, if compared to conflicts in the distant past. We weight conflict intensity according to an exponential function, using a discount factor of 0.98 for each generation of 25 years. This gives weight 1 to events in years 1850-1875, weight 0.98 to years 1825-1845, weight 0.96 to 1800-1824, and so on. An event in years 1000-1024 has half the weight of an event at the time of the Italian unification (0.50). Also this measure of conflict intensity exhibits a latitudinal ordering. On the other hand, if we adopt a lower discount factor (0.95), Faenza and Crotone become really close, and Crotone is slightly ahead (Table A-8). When we consider only defensive warfares and discard events of offensive warfare and army stationing, conflict intensity is still ordered according to latitude (Table A-9).

Table A-8: All violent conflicts in four Italian cities: Years 1000-1849.

| | Ragusa | Crotone | Faenza | Cuneo |
|--|--------|---------|--------|-------|
| Sum of all conflicts | 1 | 5 | 8 | 12 |
| Average number of conflicts per generation | 0.03 | 0.15 | 0.23 | 0.44 |
| Discounted sum of conflicts (0.98 factor) | 0.67 | 3.98 | 5.27 | 8.99 |
| Discounted sum of conflicts (0.95 factor) | 0.36 | 2.86 | 2.83 | 5.91 |

Notes: The table includes offensive warfare, defensive warfare, and army stationing as defined in the text. Average computed for years 1000-1849 for Ragusa, Crotone, Faenza and for years 1175-1849 for Cuneo. When considering 25 years the span of one generation, we have 28 generations for Cuneo, and 35 for the three other cities.

Our study provides unique evidence that experimental measures of cooperation correlate with the frequency of conflicts that has characterized the history of the four cities involved in our study. This result is in line with the hypothesis that ancestral experiences of conflict permanently shape a society’s norm of cooperation.

Table A-9: Defensive warfares in four Italian cities: Years 1000-1849.

| | Ragusa | Crotone | Faenza | Cuneo |
|--|--------|---------|--------|-------|
| Sum of defensive conflicts | 1 | 3 | 5 | 11 |
| Average number of conflicts per generation | 0.03 | 0.09 | 0.15 | 0.41 |
| Discounted sum of conflicts (0.98 factor) | 0.67 | 2.45 | 3.27 | 8.25 |
| Discounted sum of conflicts (0.95 factor) | 0.36 | 1.86 | 1.74 | 5.45 |

Notes: Average computed for years 1000-1849 for Ragusa, Crotone, Faenza and for years 1175-1849 for Cuneo. When considering 25 years the span of one generation, we have 28 generations for Cuneo, and 35 for the three other cities.

Table A-10: Count of external violent conflicts.

| time interval | Ragusa | Crotone | Faenza | Cuneo |
|---------------|--------|---------|--------|-------|
| 1000-1024 | | | | |
| 1025-1049 | | | | |
| 1050-1074 | | | | |
| 1075-1099 | | | | |
| 1100-1124 | | | 1d | |
| 1125-1149 | | | | |
| 1150-1174 | | | | |
| 1175-1199 | | | | |
| 1200-1224 | | | 1 | |
| 1225-1249 | | | | |
| 1250-1274 | | | | 1d |
| 1275-1299 | | | 1d | |
| 1300-1324 | | | | |
| 1325-1349 | | | | |
| 1350-1374 | 1d | | 1d + 1 | 1d |
| 1375-1399 | | | | 2d |
| 1400-1424 | | | 1d | 1d |
| 1425-1449 | | 1d | | |
| 1450-1474 | | | | |
| 1475-1499 | | | 1d + 1 | 1 |
| 1500-1524 | | 1 | | 1d |
| 1525-1549 | | 1 | | 1d |
| 1550-1574 | | 1d | | 1d |
| 1575-1599 | | | | |
| 1600-1624 | | | | |
| 1625-1649 | | | | 1d |
| 1650-1674 | | | | |
| 1675-1699 | | | | 1d |
| 1700-1724 | | | | |
| 1725-1749 | | | | 1d |
| 1750-1774 | | | | |
| 1775-1799 | | | | |
| 1800-1824 | | 1d | | |
| 1824-1849 | | | | |
| 1850-1861 | | | | |

Notes: We denote with “d” the *defensive warfare*; other conflicts are either *offensive warfare* or *army stationing*.



Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

Strada Maggiore 45
40125 Bologna - Italy
Tel. +39 051 2092604
Fax +39 051 2092664
<http://www.dse.unibo.it>