INSTITUTE
OF ECONOMICS

Scuola Superiore
Sant'Anna

LEM | Laboratory of Economics and Management

Institute of Economics
Scuola Superiore Sant'Anna

Piazza Martiri della Libertà, 33 - 56127 Pisa, Italy
ph. +39 050 88.33.43
institute.economics@sssup.it

LEM
WORKING PAPER SERIES

When the two ends meet: an experiment on cooperation across the Italian North-South divide

Pietro Battiston [a]
Simona Gamba [b]

[a] Istituto di Economia, Scuola Superiore Sant'Anna, Pisa, Italy
[b] FBK-IRVAPP, Trento, Italy

2016/41 **December 2016**
ISSN(ONLINE) 2284-0400

When the two ends meet: an experiment on cooperation across the Italian North-South divide

Pietro Battiston[*], Simona Gamba[†]

December 2016

Abstract

We study the behavior of individuals coming from different geographic regions of Italy, in a same public good game. We confirm previous findings according to which, faced with the same incentives and experimental conditions, Southern citizens exhibit a lower propensity to cooperate than Northern ones. This difference is mainly explained by a gap in the impact of coordination devices available to participants, as we show by manipulating them. Most importantly, when subjects with different geographic origins are teamed up together, their contributions decrease with respect to homogeneous groups, again because of a reduced effect of coordination devices. These findings reinforce the interpretation of the Italian South-North divide as related to trust, prejudice and a consequent path-dependence in levels of social capital, rather than due to the mere effect of differences in institutions and economic opportunities.

Keywords: public good, cooperation, social capital, cultural differences, laboratory experiment.
JEL classification: A13, C71, C92, H41.

[*]Istituto di Economia, Scuola di Studi Superiori Sant'Anna, Piazza Martiri della Libertà, 33, 56127 Pisa, Italy. me@pietrobattiston.it

[†]FBK-IRVAPP, via Santa Croce, 77, 38122 Trento, Italy. gamba.simona@gmail.com

1 Introduction

Almost since the dawn of experimental economics, researchers have looked with ever growing interest at what experiments ran in various geographic locations could reveal concerning the specific characteristics of different cultures and societies. This is particularly true for experiments focusing on traits related to *social capital*, such as trust (Croson and Buchan 1999), cooperation (Cason et al. 2002) and fairness (Oosterbeek et al. 2004), given the fundamental importance that social capital bears in explaining differences across regions of the world in terms of institutional organization and economic outcomes (Knack and Keefer 1997; Buonanno et al. 2009; Hoyman et al. 2016).

In Italy, profound internal differences in terms of social capital and cooperation have since long been identified as one of the causes of the North-South economic divide that has characterized the country since its unification and has widened in the last decades (Helliwell and Putnam 1995; Leonardi 1995; Guiso et al. 2004). Although the historical origins of the gap remain a matter of debate, empirical evidence on the existence of the divide itself is overwhelming, based not only on economic indicators (GDP per capita, unemployment rate, internal migrations), but also on measures of quality of institutions (timeliness of budgets, legislative innovation, citizen satisfaction) and on individual level indicators (frequency of blood donations, number of associations, voters turnout at elections, newspaper readership).[1] Regional disparities have been often explained on the ground of differences in economic opportunities and quality of institutions, but a stream of literature, which can be reconduced to the seminal work of Banfield (1967), has focused instead on the individual determinants of the propensity to cooperate. Ichino and Maggi (2000), for instance, exploited the phenomenon of on-the-job movers inside a large Italian bank to compare individuals facing the same incentives but having different geographic backgrounds. Bigoni, Bortolotti, Casari, Gambetta, and Pancotto (2016) run a "laboratory-in-the-field" experiment in two cities located in the North and two located in the South of Italy, with experimental subjects being presented the exact same incentives and experimental conditions. Their results confirm that observed disparities in behavior cannot be explained just by differences in the economic context,

1. The empirical literature on the *"Questione meridionale"*, i.e. the North-South gap, is vast: see Helliwell and Putnam (1995), Ichino and Maggi (2000), Felice (2013), and Bigoni, Bortolotti, Casari, Gambetta, and Pancotto (2016) for a more comprehensive view.

but rather that they are *"likely to derive from persistent differences in social norms"*. Far from being a purely Italian problem, differences in social capital among regions of a same nation are studied in other countries of the world: see, for instance, the experimental literature on West and East Germany (Ockenfels and Weimann 1999; Brosig-Koch et al. 2011).

In this paper, through a "field-in-the-laboratory" experiment, we study the propensity of individuals from both the North and South of Italy to contribute in a same public good game. By manipulating both the composition of the groups taking part in our experiment and the available coordination opportunities, we improve upon the existing literature by exploring the effect of such opportunities on contributions, and by shedding light on what determines the differences in the level of contributions between Northern and Southern citizens. A crucial novelty of our study, compared to the experimental literature mentioned so far, is then the ability to observe the interaction *between* individuals characterized by different geographic backgrounds. Aside from the experimental literature, a growing stream of research focuses on the comparison of migrants and on-the-job movers to local populations (Ichino and Maggi 2000; Gibson et al. 2014; Algan et al. 2016); however, our setting is unique in the fact that it abstracts the effect of migration and integration (or segregation). Our subjects were living in different cities *at the time of the experiment*, they moved to the location were the experiment took place only for few days, and except for the geographic background they shared similar characteristics (such as age and education).

Our results confirm differences in reactions to identical incentives: specifically, a lower level of cooperation characterizing Southern citizens. In our experiment, this is mostly explained by a gap in the impact of coordination opportunities. In a closely related study, Bigoni, Bortolotti, Casari, and Gambetta (2016) highlight the fact that individuals from the South expect lower contributions from their peers, but feature a *higher* level of conditional cooperation. These evidences emphasize the need to consider the behavioral aspects of the North-South divide in Italy: even though behavioral traits might have certainly evolved as an adaptation to institutional characteristics (as already suggested by Putnam et al. 1993), it can be misleading to expect that changing such characteristics will have an immediate positive effect on cooperation levels. Such intuition can help explaining the failure of past measures adopted by policymakers in order to close the North-South gap, and must be taken into account when planning further actions in this direction (a current example being represented by policies aimed at improv-

ing the competitiveness of universities located in the South of Italy, e.g. by attracting winners of ERC grants).

Most importantly, we observe that individuals in mixed groups cooperate less compared to those in homogeneous groups, once they get to know the composition of their group. Although anonimity in public good games was already studied by Andreoni and Petrie (2004), among others, the novelty of our study lies in the analysis of within-subject behavior *across* rounds of a same game. These aspects are crucial for the debate on economic inequalities across geographic regions, since they highlight the possible role of prejudice and, more in general, lack of integration. In areas where social capital is scarce, economic development may be also hindered by the relatively difficult interactions with other regions: in absence of measures overcoming regional disparities, these might spontaneously deteriorate over time.

The following section describes the characteristics and design of the experiment, Section 3 presents our hypotheses, Section 4 the results and Section 5 concludes.

2 Experimental design

The experiment was ran on October 3, 2015, in Volterra (Italy), and involved students from 13 schools located in 7 different cities, part of 5 different Italian regions. The context was a project organized by Sant'Anna School of Advanced Studies: all students were in their last year of school, hence between 17 and 18 years of age. Our subjects share a medium social background which makes them presumably more representative of the Italian population than samples of university students typically involved in experiments. Specifically, our subjects have a mother not holding a university degree, and a track record of relatively good grades.[2] Most importantly, the *geographic* representativeness of our sample is an exception in the experimental literature, in particular when considering that experimental subjects were, *at the time of the experiment*, living in the 7 cities of origin.[3] In total, the experi-

2. The general scope of the project was related to curriculum counseling: the literature on inter-generational transmission on education points at the mother's level of education as particularly relevant (Black and Salvanes 2005; Pronzato 2012).

3. Bigoni, Bortolotti, Casari, and Gambetta (2016) for instance look at a sample of students of the University of Bologna originating from a large number of cities, located in the North and in the South of Italy.

ment involved 78 subjects (49 females and 29 males): 42 subjects came from schools in the South of Italy and 36 from schools in the Center-North.[4]

Four experimental sessions were ran, each involving 19 or 20 participants. In each session, participants were regrouped into four groups. Of such groups, one was composed only by students coming from the schools in the South, one only by students coming from the schools in the Center-North, while the other two had mixed composition: *"being member of a homogeneous group"* is our main treatment variable. The groups were formed ex ante randomly, with the condition that no two students from the same school would end up in the same group, and that the sizes of the groups were as homogeneous as possible, given the requirement defined above.[5] For the later interpretation of the results, it is worth mentioning that participants had been involved in group activities for the last two days, and hence knew each other at least superficially, while it is rather improbable that subjects coming from different schools knew each other before then. It is also worth mentioning that to Italian speakers, Southern and Northern accents are very easily distinguishable. Thus, it is highly plausible that, at the time our experiment was run, participants were broadly aware of each other's origin. On the other hand, at no time during the experiment was any reference to geographic origin, or to the North-South divide, made.

In each session, six rounds of a linear public good game were played. At each round, each participant was given four playing cards, which only she or he could observe. Two of them were red, and were worth one point each; the other two were black, and were worth zero points. Two cards were then

4. The seven cities involved in the experiment were Cagliari, Napoli, Palermo, Partinico for the South, and Massa, Milano, Prato for the Center-North. Participants from the Center are pooled with those from the North in light of the characteristics of their cities of origin, both located in Toscana. Bigoni, Bortolotti, Casari, Gambetta, and Pancotto (2016), in their selection procedure, classify Toscana in the North based on its latitude. Such choice is reinforced by a look at socioeconomic variables they adopt as proxies for social capital: according to the data they provide, when compared to the average for Northern Italy, Toscana has higher association density (68.44 per 100,000 inhabitants vs. 36.57, South is at 23.97) and electoral participation (86.7% vs. 86.0%, South is at 70.2%), while it is close to the North average for blood donations (42.5 every 1000 inhabitants vs. 47.9, South is at 23.5).

5. All groups were designed to have five or six members, but five groups out of sixteen had only four members due to absences. No group had more than three members from a same city (the algorithm used for creating the groups is described in detail in Appendix A).

collected, covered, from each participant, who could therefore secretly decide to give zero, one or two points (red cards). The total amount of points collected within each group was multiplied by 2 and subdivided between participants of that group. Such points were then added to each participant's "private earnings" - the number of red cards she or he had decided to keep - so that the total gain for an individual i in a given round t would be:

$$\pi_{i,t} = \underbrace{2 - x_{i,t}}_{\text{Private earnings}} + \underbrace{\frac{2}{\mathcal{N}} \sum_{j=1}^{\mathcal{N}} x_{j,t}}_{\text{Public earnings}}$$

with $x_{i,t}$ being the individual contribution to the public good, $\pi_{i,t}$ the individual earnings and $\mathcal{N}$ the group size.

Each participant then received back her two cards,[6] still hidden from the view of others participants, and the next round began. Earnings were summed across all rounds,[7] and prizes were assigned, in each session, to the three players which had cumulated most points after the six rounds. These rules were explained in advance to participants, who were invited to ask questions in case any aspect was unclear.

The six rounds differed in the amount of information and coordination opportunities available to participants (see Figure 1).

- Initially, the students were sitting in circle around a room, in an order, previously determined by the experimenters, satisfying the condition that neighbors were not in the same group. They knew that they were subdivided in four groups of roughly equal size, and that such groups would have been kept unchanged for the six rounds of the game, but they did not know who their groupmates were.

- After round 2, the names of members of each group were made public, ensuring, by asking them to raise their hands, that participants of each

6. The mapping between cards and participants was fixed since the beginning, allowing the experimenters both to record private earnings, and to return to each player the contributed cards after each round. For practicality, each participant was assigned four cards with the same number or face, two from a black suit and two from a red suit: for instance, "10 of clubs and diamonds".

7. By allowing for potential carry-over effects, we are able to study intertemporal group dynamics more in depth, and in a more natural setting. This design choice is consistent with studies on repeated public good games against which we compare our results (Andreoni 1988; Andreoni and Petrie 2004).

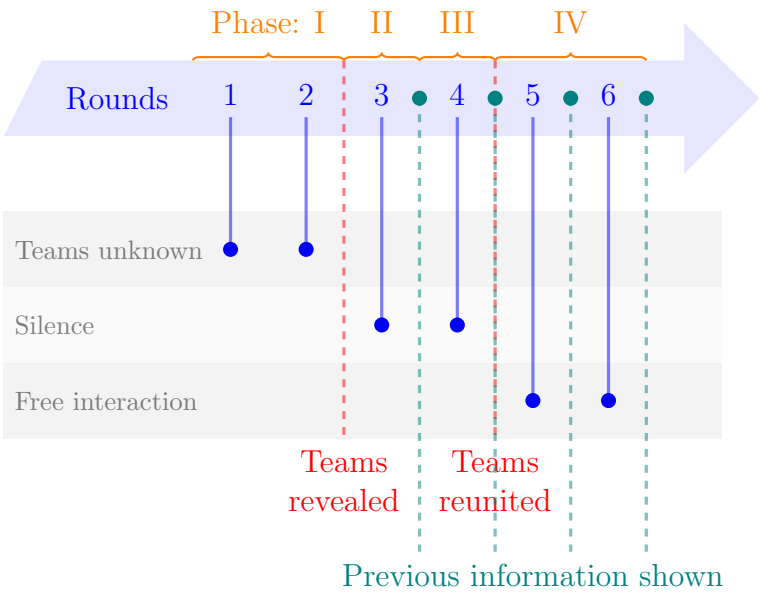


Figure 1: Timeline of the experiment

group had identified each other visually. Participants were then asked not to communicate in any way among them, until further notice.

- After round 4, participants were instructed to sit together with their groupmates, with each group in a different corner of the room, and were given two minutes to discuss among them. The same happened after round 5.

Moreover, after each round starting from the third, information about past contributions was released to participants in two ways: individual contributions from the previous round were read aloud but anonymously, i.e. by referring to the cards owned by each individual rather than to her name, and past results for each group were shown graphically to participants (for an example, see Figure 3 in Appendix B).

These changes in design allow us to investigate the issues of information, anonymity (particularly important in the Italian context, given that the North-South gap has been frequently related to trust and to cultural distance) and coordination.

3 Hypotheses

We want to first test whether, in line with the existing literature, the propensity to contribute to the public good is related to the geographic origin. To this aim, we compare the average contribution of individuals from the North and from the South, by testing the following hypothesis:

$$\mathcal{H}_0 : \{\bar{x}_i\}_{N_i=0} = \{\bar{x}_i\}_{N_i=1}, \tag{Hn}$$

where the bar denotes the individual average over all periods, and N_i is a dummy variable taking value 1 if subject i is from the North, and 0 otherwise.

We then analyze the effect of design changes on contribution levels: for this, we need to consider separately the different rounds of the game. Rounds 1 and 2 present the same information setting, and will be analyzed together; the same can be said for rounds 5 and 6. Instead, rounds 3 and 4 differ, since before round 4 (but not before round 3) individuals were given aggregated information on their group's contributions history (and they knew that this information would be communicated after each of the following rounds). Hence, we will refer to rounds 1 and 2 as "phase I", round 3 as "phase II", round 4 as "phase III" and rounds 5 and 6 as "phase IV": each phase coincides with a different level of information/coordination opportunities.

In order to test the effect of such changes of setting, we start by estimating the following model:

$$\begin{aligned} x_{i,t} =& \alpha^f F_i + \alpha_I T_{t,I} + \alpha_{II} T_{t,II} + \alpha_{III} T_{t,III} + \alpha_{IV} T_{t,IV} + \epsilon_{i,t} \\ =& \alpha^f F_i + \sum_{P=I}^{IV} T_{t,P} \alpha_P + \epsilon_{i,t}, \end{aligned} \tag{1}$$

where each phase dummy $T_{t,P}$ takes value 1 if t is in phase P,[8] and F_i takes value 1 if individual i is a female.

In principle, a positive value for phases coefficients could be a spurious consequence of learning effects. However, this is categorically ruled out by a stylized fact coming from the literature on repeated public goods games: when subjects are informed about the length of the game, *"provision of the*

8. We insert a dummy for *each* phase, including the first: coherently, we do not insert a constant in the model. This choice clearly does not affect the results (we will look at comparisons between coefficients α_P rather than at their absolute value), and it greatly simplifies our equations.

public good 'decays' toward the free riding level with each repetition" (Andreoni 1988). Hence, any significant increase in contributions across phases can be considered as (a lower bound to) the effect of the changes in design.

To verify whether the mere identification of groupmates affects contributions, we start by testing the following hypothesis:[9]

$$H_0 : \alpha_{II} = \alpha_I. \tag{HcII}$$

We then test the effect of information concerning previous contributions:

$$H_0 : \alpha_{III} = \alpha_{II} \tag{HcIII}$$

and the effect of communication:

$$H_0 : \alpha_{IV} = \alpha_{III}. \tag{HcIV}$$

Concerning the treatment variable "belonging to a homogeneous group", denoted as HOM_i, we first test whether members of homogeneous groups exhibit a higher propensity to contribute to the public good, in a way analogous to hypothesis (Hn):[10]

$$\mathcal{H}_0 : \{\bar{x}_i\}_{HOM_i=0} = \{\bar{x}_i\}_{HOM_i=1}. \tag{Hh}$$

We then analyze the treatment effect across the different settings by interacting it with phase dummies:

$$x_{i,t} = \beta^f F_i + \sum_{P=I}^{IV} T_{t,P}(\beta_P + \beta_P^h HOM_i) + \epsilon_{i,t}. \tag{2}$$

This is the central equation for estimating the effect of group composition. Importantly, the interpretation of its coefficients is only meaningful from phase II onwards, since during phase I subjects do not have any information on their group, and so there cannot be a treatment effect.

Focusing on the transition from phase I to phase II, we can first check the effect of identification on contributions, restricting to mixed groups:

$$\mathcal{H}_0 : \beta_{II} = \beta_I, \tag{HmII}$$

9. The "c" in the hypothesis label refers to the *context* of the decision.

10. We will exclude observations from phase I, when the groups composition is still unknown to participants.

and to homogeneous groups:

$$\mathcal{H}_0 : \beta_{II} + \beta^h_{II} = \beta_I + \beta^h_I. \tag{HhII}$$

Second, we check whether the identification of group participants represents an advantage for homogeneous groups with respect to mixed groups. This can be done in two ways. One is to check whether individuals in homogeneous groups contribute more than individuals in mixed groups in phase II:

$$\mathcal{H}_0 : \beta^h_{II} = 0. \tag{HdII}$$

The other is to test the difference between the two groups across the two phases, following a Difference-in-Differences approach:[11]

$$\mathcal{H}_0 : \beta^h_{II} - \beta^h_I = \beta_{II} - \beta_I. \tag{HddII}$$

The four hypotheses just presented focus on the transition from phase I to phase II, looking respectively at the mixed groups (HmII), the homogeneous groups (HhII), the difference between the two (HdII) and the Difference-in-Differences (HddII). The same approach can be applied to the other two changes in the information setting. We will hence have a similar set of hypotheses focusing on the transition from phase II to phase III (i.e. on the effect of information concerning past contributions), which follows:

$$\mathcal{H}_0 : \beta_{III} = \beta_{II} \tag{HmIII}$$
$$\mathcal{H}_0 : \beta_{III} + \beta^h_{III} = \beta_{II} + \beta^h_{II} \tag{HhIII}$$
$$\mathcal{H}_0 : \beta^h_{III} = 0 \tag{HdIII}$$
$$\mathcal{H}_0 : \beta^h_{III} - \beta^h_{II} = \beta_{III} - \beta_{II}. \tag{HddIII}$$

Finally, an analogous set of hypotheses focuses on the transition from phase III to phase IV (i.e. on the effect of communication among groupmates):

$$\mathcal{H}_0 : \beta_{IV} = \beta_{III} \tag{HmIV}$$
$$\mathcal{H}_0 : \beta_{IV} + \beta^h_{IV} = \beta_{III} + \beta^h_{III} \tag{HhIV}$$
$$\mathcal{H}_0 : \beta^h_{IV} = 0 \tag{HdIV}$$
$$\mathcal{H}_0 : \beta^h_{IV} - \beta^h_{III} = \beta_{IV} - \beta_{III}. \tag{HddIV}$$

11. The two approaches are interchangeable at this time, given that the two groups are virtually indistinguishable during phase I. This is not the case for the following phases.

Equation (1) considers the effect of different information settings without disaggregating on the geographic origin of subjects. We instead verify whether the ability to exploit coordination possibilities is related to the geographic origin by interacting the phase dummies with the N_i dummy (indicating whether subject i is from the North):

$$x_{i,t} = \gamma^f F_i + \sum_{P=I}^{IV} T_{t,P}(\gamma_P + \gamma_P^n N_i) + \epsilon_{i,t}. \tag{3}$$

Equation (3) allows us to test the hypothesis $\mathcal{H}_0 : \gamma_P^n = 0$ for each phase $P = I, II, III, IV$, answering to the question of whether the effect of increases in coordination opportunities relates to the geographic origin.

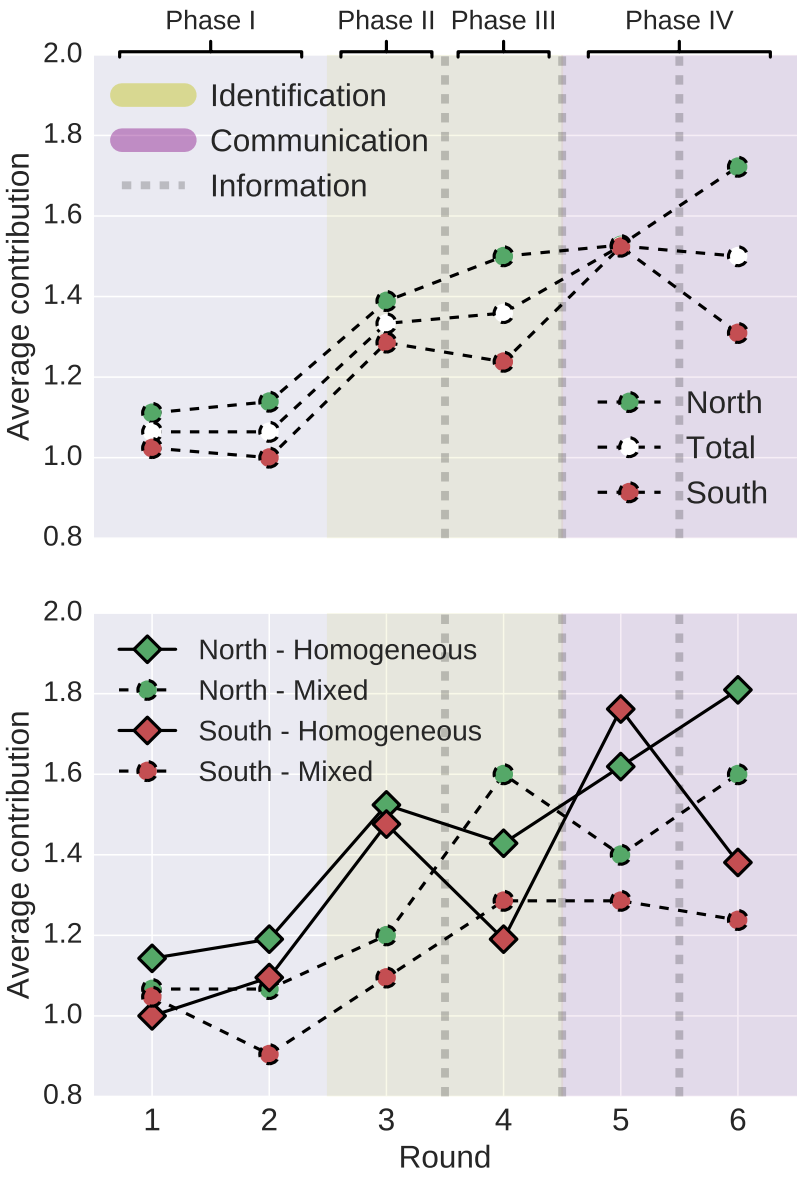
4 Results

The average contribution across all sessions and rounds was 1.308. Figure 2 (top, white dots) plots average contributions in each round: the effect of changes in coordination possibilities is evident between phases I and II, and between phases III and IV. Instead, no evident change can be detected between phases II and III, possibly because the effect of information about group contributions can depend on such contributions (e.g. because of conditional cooperation or, conversely, of incentives to free ride). In fact, the disaggregation by geographic origin shows that, when moving to phase III, the behavior differs between individuals from the North and from the South (red and green dots), while the disaggregation by treatment status (Figure 2, bottom) shows an even more pronounced difference: contributions increase in mixed groups and decrease in homogeneous ones.

4.1 Treatment effect

We now systematically analyze hypotheses formulated in Section 3, starting from the treatment effect (Hh). As already mentioned, we test it on data from phases II to IV: in phase I, participants did not know the composition of their group (we analyze this phase in Section 4.4 as a robustness test on the randomization process). Since different phases clearly do not provide independent information, we run a a Mann-Whitney test on individual averages over rounds. We find that $\{\bar{x}_i\}_{HOM_i=0} < \{\bar{x}_i\}_{HOM_i=1}$ (participants in

Figure 2: Average contributions per round



Note: Top: average contributions by geographic origin. Bottom: average contributions by geographic origin and group composition.

homogenous groups contribute more), and that the difference is significant ($p = 0.040$).

> **Result 1** *Groups composed by members sharing the same geographic origin contribute to the public good more than mixed groups.*

The already mentioned increase in contributions across phases, which is evident in Figure 2, is *per se* a nontrivial finding, given the decay in contributions over time consistently observed by the experimental literature (Andreoni 1988), even in absence of anonymity (Andreoni and Petrie 2004). Hence, we can infer that changes of setting have an effect in increasing contributions: we now proceed to a more formal analysis of such effect. Estimated coefficients for equations (1), (2) and (3) are presented in Table 1.

What follows is the summary of results concerning the identification of groupmates, which happens in phase II.

- From Equation (1):
  - (HcII): $\alpha_{II} > \alpha_I$ ($p = 0.015$)
- From Equation (2):
  - (HmII): $\beta_{II} = \beta_I$ not rejected ($p = 0.450$)
  - (HhII): $\beta_{II} + \beta^h_{II} > \beta_I + \beta^h_I$ ($p = 0.001$)
  - (HdII): $\beta^h_{II} > 0$ ($p = 0.048$)
  - (HddII): $\beta^h_{II} - \beta^h_I > \beta_{II} - \beta_I$ ($p = 0.000$)

Identification of group members has a positive and significant effect on contributions (HcII): this is driven by subjects in homogeneous groups (HhII), who contribute significantly more than the others (HdII), while no significant change can be found for mixed groups (HmII). Indeed, the Difference-in-Differences is also significant (HddII), confirming that the ability to distinguish groupmates affects subjects in homogeneous groups more than subjects in mixed groups. Notice that, in mixed groups, the identification of groupmates does not reveal, at an aggregate level, new information concerning the group composition: on average, the group has the same share of participants from the South and from the North as the entire session. This can explain at least in part why the transition from phase I to phase II produces no significant effects in mixed groups.

Table 1: Main results

		Eq. (1) (α)	Eq. (2) (β)	Eq. (3) (γ)
	Female	0.133	0.082	0.149
		(0.089)	(0.081)	(0.086)
Phase	I	0.981***	0.977***	0.913***
		(0.095)	(0.118)	(0.097)
	II	1.250***	0.102***	1.186***
		(0.112)	(0.144)	(0.123)
	III	1.275***	1.380***	1.139***
		(0.098)	(0.103)	(0.107)
	IV	1.429***	1.324***	1.317***
		(0.094)	(0.140)	(0.111)
Phase interacted with treatment	h,I		0.065	
			(0.117)	
	h,II		0.333**	
			(0.155)	
	h,III		-0.135	
			(0.148)	
	h,IV		0.254	
			(0.157)	
Phase interacted with North	n,I			0.126
				(0.081)
	n,II			0.116
				(0.152)
	n,III			0.274*
				(0.131)
	n,IV			0.221**
				(0.102)
	N	468	468	468

Note: The dependent variable in each model is $x_{i,t}$. Row labels indicate coefficients subscripts: phase dummies are in the first block, phase-treatment interactions in the second block, phase-origin interactions in the third block. E.g. the bottom right cell contains the estimate of γ^n_{IV}, relative to the interaction of dummy variables $T_{t,IV}$ (fourth phase) and N_i (North) in Equation 3. Group-level clustered standard errors in parentheses.

***p< 0.01, **p< 0.05, *p< 0.10.

Result 2 *Identification of groupmates significantly increases contributions only in groups composed by members sharing the same geographic origin.*

We now present the summary of results concerning the transition from phase II to phase III.

- From Equation (1):
  - (HcIII): $\alpha_{III} = \alpha_{II}$ not rejected ($p = 0.818$)
- From Equation (2):
  - (HmIII): $\beta_{III} > \beta_{II}$ ($p = 0.075$)
  - (HhIII): $\beta_{III} + \beta^h_{III} = \beta_{II} + \beta^h_{II}$ not rejected ($p = 0.126$)
  - (HdIII): $\beta^h_{III} = 0$ not rejected ($p = 0.374$)
  - (HddIII): $\beta^h_{III} - \beta^h_{II} < \beta_{III} - \beta_{II}$ ($p = 0.000$)

As already suggested by Figure 2 (top, white dots), observing past group performance does not significantly affect *average* contributions (HcIII). In the bottom plot, we can observe a sort of rebound effect: contributions in round 4 decrease for homogeneous groups, which were the best performers in round 3, while the opposite stands for mixed groups. Indeed, the increase in contributions in mixed groups is significant (HmIII) and, while the decrease in contributions in homogeneous groups is not (HhIII), the difference between the two is (HddIII).[12]

We finally verify how the possibility to communicate (phase IV) affects contributions.

- From Equation (1):
  - (HcIV): $\alpha_{IV} = \alpha_{III}$ not rejected ($p = 0.240$)
- From Equation (2):

12. Individuals observed (aggregated) information concerning past contributions (as in Figure 3, Appendix B), and hence could be affected by issues of conditional coperation. Moreover, since they observed the performance of *all* groups and they knew that prizes would go to the three best performers in the session, they could decide to increase free riding in order to capitalize on a good group standing, or to limit it in order to catch up.

- (HmIV): $\beta_{IV} = \beta_{III}$ not rejected ($p = 0.796$)
- (HhIV): $\beta_{IV} + \beta^h_{IV} > \beta_{III} + \beta^h_{III}$ ($p = 0.012$)
- (HdIV): $\beta^h_{IV} = 0$ ($p = 0.127$)
- (HddIV): $\beta^h_{IV} - \beta^h_{III} < \beta_{IV} - \beta_{III}$ ($p = 0.000$)

Comparing phase IV with phase III, we do not find a significant increase in average contributions (HcIV). Indeed, we do find a positive variation for homogeneous groups (HhIV), but not for mixed ones (HmIV), and we observe a significant difference across phases between the two categories (HddIV). However, when we do not consider the comparison with the previous phase, the difference between the two categories is not significant (HdIV), and this forbids us from drawing unambiguous conclusions on the effect of communication (that is, compared to a mere rebound effect).

4.2 Effect of geographic origin

The average contribution of subjects from the North is 1.398, while it is 1.230 for subjects from the South. By testing (Hn), we ascertain whether this difference is significant. A Mann-Whitney test on average contributions rejects the null hypothesis, evidencing that $\{\bar{x}_i\}_{N_i=1} > \{\bar{x}_i\}_{N_i=0}$ ($p = 0.010$). While this approach clearly evidences that subjects from the North contribute more than subjects from the South to the public good, the interpretation of this result is nontrivial. Indeed, being in a "North-only group" can have a different effect than being in a "South-only group" on the propensity to contribute, and to the extent to which this is true, the result just presented can be affected by the treatment. In order to isolate the individual-level geographic effect, we hence run the same test restricting the sample to mixed groups (members of which are not affected by differences in treatment), again rejecting the null hypothesis ($p = 0.047$). This evidence is in line with the available experimental literature on the North-South gap in Italy (Bigoni, Bortolotti, Casari, Gambetta, and Pancotto 2016).

Result 3 *Subjects from the North contribute to the public good more than subjects from the South.*

Concerning the analysis of the geographic effect across phases (Equation 3 - see Table 1 for estimates), what emerges is that γ^n_P is (positive and) significantly different from zero in the last two phases only ($p = 0.141, 0.459, 0.054, 0.048$, respectively). We can hence state the following:

Result 4 *The higher level of contributions of subjects from the North is explained by a stronger reaction to the introduction of coordination opportunities rather than by a higher propensity to contribute since the first rounds.*

Notice that interaction coefficients of Equation 3 could be affected by a potentially asymmetric treatment effect (being part of a South-/North-only group). Disaggregating further the analysis, by combining the two aspects of group homogeneity and geographic origin, would allow us to tackle this issue and also to verify whether the treatment effect itself (Result 1) is to be attributed in larger part to North-only groups, or to South-only ones. We do so in Equation 5, presented in Appendix C; however, the increase in the number of regressors can by itself justify the mostly non-significant results.[13] The two significant interaction coefficients (evidencing higher contributions, in mixed groups, of Northern students) are in line with Result 3, and the fact that they refer to the last phases is a confirmation of Result 4.

4.3 Contributions and gender

In the literature on public good games, some evidence has been provided (Nowell and Tinkler 1994) of a higher propensity to contribute on behalf of females. Estimates reported above have been obtained controlling for a potential gender effect; in the present section, we verify whether females and males have a different propensity to contribute to the public good, and whether they react differently to the treatment and/or to the changes in information settings.

We can test the presence of a gender effect by running a Mann-Whitney test on average contributions of females versus males: the result is not significant ($p = 0.106$). To avoid possible confounding effects related to the treatment, we run the same test restricting to mixed groups (like we did for hypothesis (Hn)): the result is again not significant ($p = 0.180$).

13. For example, since there is a significant effect of being assigned to a homogeneous group (results 1 and 2) at the aggregate level, then this must also hold in either North-only or South-only groups (or both). Given that our two geographic subsamples are of roughly equal size, re-running the same experiment on a double sized sample might allow us to analyze each of the two subsamples with a power equivalent to the tests presented so far, and hence to shed additional light on the combined effect of treatment and geographic origin.

For what concerns the gender component *of the treatment effect*, we can estimate the following equation (analogous to Equation 5 in Appendix C, but with geographic origin replaced by gender):

$$x_{i,t} = \sum_{P=I}^{IV} T_{t,P}(\delta_P + \delta_P^h HOM_i + F_{i,P}(\delta_P^f + \delta_P^{hf} HOM_i)) + \epsilon_{i,t}. \qquad (4)$$

See Table 4 in Appendix C for estimation results. By testing $\mathcal{H}_0 : \delta_P^f + \delta_P^{hf} = 0$ for $P = I, II, III, IV$, we verify whether in homogeneous groups females behave differently than males. Results are never significant ($p = 0.484, 0.739, 0.604, 0.514$): that is, we find no evidence of a gender difference in the effect of the treatment.

4.4 Robustness

Result 1, concerning the effect of the treatment, is supported by a Mann-Whitney test, which does not require any distributional assumption. It can be alternatively obtained by testing the joint significance of phase-treatment interaction dummies in Equation 2: we do so through a Wald test, and again reject the null hypothesis of no difference ($p = 0.078$). Analogously, Result 3, concerning the effect of geographic origin, is confirmed by looking at the joint significance of phase-origin interaction dummies in Equation 3 ($p = 0.054$).

As already mentioned in Section 3, Equation 2 is not expected to yield interesting insights concerning phase I: at that time, subjects did not know who their groupmates were, and hence their contribution could not be affected by being in a homogeneous or mixed group. If homogeneous and mixed group members had differed in their contribution levels already before the group composition was made public, this would have represented an alarming signal of ex ante differences among the two samples. However, this is not the case, and the null hypothesis that $\beta_I^h > 0$ cannot be rejected ($p = 0.586$).

Female participants outnumbered male participants in all sessions. We both allowed for a gender effect in our estimates, and explicitly looked at a gender component of the treatment effect in Section 4.3, without finding any. This said, as discussed in Appendix B, in one specific case (session 2), the randomization algorithm resulted in a significantly unbalanced composition of homogeneous groups, which included no male participants. We hence re-estimate hypothesis (Hh) excluding such groups from the analysis, and still find a significantly positive treatment effect ($p = 0.079$).

5 Conclusions

We run a public good game in which participating subjects come from different Italian cities. Compared to the existing experimental literature on differences in social capital across geographic regions, we observe the interaction of subjects who are temporarily abstracted from their city and, most importantly, who have different geographic backgrounds. By manipulating the composition of groups, we compare the ability to cooperate of those constituted only by individuals sharing the same origin, on the one side, and mixed ones, on the other. At the same time, we explore to which extent the identification of groupmates and communication with them can lead an individual to contribute more to the public good, and we compare the contribution patterns of individuals with different geographic backgrounds.

We find that groups composed by subjects both from the North and from the South of Italy perform significantly worse than homogeneous groups: this is mainly due to the different extent to which they exploit implicit (identification) and explicit (communication) coordination devices. As already reported in the literature, individuals from the North contribute more than individuals from the South, and we additionally show that this happens even *within the same group*. This is explained by a different reaction to information concerning past contributions, and to the possibility to communicate. Instead, the two groups do not differ significantly in the first phase of the experiment: hence, there is no evidence of a difference in the ex ante propensity to contribute. In general, the effect of coordination devices is strongly positive, and their introduction more than counteracts the expected decay of contributions over time. Finally, while females tend to contribute more than males, the difference is not significant, neither in average contributions nor in their reaction to the treatment.

Our results shed new light on the problem of the North-South divide in Italy. They reinforce (Result 3) the conclusions of Bigoni, Bortolotti, Casari, Gambetta, and Pancotto (2016) that the gap *"appears to lie in the ability to cooperate"*. In particular, we show (Result 4) that the effectiveness of coordination, rather than just the propensity to contribute to a common good and to trust one's peers, can make a difference. Since geographic origin is not a significant predictor of contributions *at the beginning* of the experiment (when participants act in isolation), we show that prejudices, and more in general mutual trust, play a fundamental role in the North-South divide. At the same time, our findings (Result 1, 2) highlight the difficulty in co-

operation *across* the North-South divide: such difficulty could be the cause of strong path dependence (historically, the gap in socioeconomic indicators does not seem to vanish over time, rather the converse), and further obstacle economic and social development.

Summing up, our results add skepticism towards the idea that the North-South gap in social capital can be imputed only to differences in institutions and opportunities, pointing out at the different reactions to the same incentives. Hence, such gap cannot be leveled by only focusing on institutional settings: while, in the long term, behavior can certainly react to institutions, such reactions might be too slow. Most importantly, institutions themselves are composed of citizens, and any attempt at shaping them must take this aspect into account. This said, the evidence provided can indicate to policymakers a path to follow. Like most of the Italian population, the typical participant in our experiment had relatively few occasions to enter in relation with compatriots from the other side of the peninsula: the literature on the positive effect of diversity on economic performance (Lazear 1999; Hong and Page 2001) is well aware of the problem of communication costs, and suggests that policymakers should work in the direction of integration and mutual knowledge. These should be considered among the main targets when dealing with socioeconomic differences across geographic regions, and as a viable way to increase the level of social capital in countries characterized by strong etherogeneities.

Further studies could be devoted to analizing the interaction of the treatment with the geographic origin. Indeed, we do not find a significant effect of being in a homogeneous group *conditional* on being from the North/South: additional experimental evidence might allow to achieve such level of detail. This will allow for instance to state whether the worse performance of mixed groups can be imputed more to one of the two geographically defined subpopulations, or whether individuals from the North contribute less when they are in mixed groups than when they are in North-only groups. We think that these are important issues to consider for the understanding of the North-South economic divide, and that they are an interesting venue for future research, together with the study of other brackets of the population.

References

Algan, Yann, Camille Hémet, and David D Laitin. 2016. "The Social Effects of Ethnic Diversity at the Local Level: A Natural Experiment with Exogenous Residential Allocation." *Journal of Political Economy* 124 (3): 000–000.

Andreoni, James. 1988. "Why free ride?: Strategies and learning in public goods experiments." *Journal of Public Economics* 37 (3): 291–304.

Andreoni, James, and Ragan Petrie. 2004. "Public goods experiments without confidentiality: a glimpse into fund-raising." *Journal of public Economics* 88 (7): 1605–1623.

Banfield, Edward C. 1967. *The moral basis of a backward society.* Free Press.

Bigoni, M., S. Bortolotti, M. Casari, and D. Gambetta. 2016. "At the Root of a Difference: Cooperation, Expectations and Preferences in a Divided Country." *Mimeo.*

Bigoni, M., S. Bortolotti, M. Casari, D. Gambetta, and F. Pancotto. 2016. "Amoral Familism, Social Capital, or Trust? The Behavioural Foundations of the Italian North-South Divide." *The Economic Journal* 126:1318–1341.

Black, Paul J., Sandra E. amd Devereux, and Kjell G. Salvanes. 2005. "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital." *The American Economic Review* 95 (1): 437–449.

Brosig-Koch, Jeannette, Christoph Helbach, Axel Ockenfels, and Joachim Weimann. 2011. "Still different after all these years: Solidarity behavior in East and West Germany." *Journal of Public Economics* 95 (11): 1373–1376.

Buonanno, Paolo, Daniel Montolio, and Paolo Vanin. 2009. "Does social capital reduce crime?" *Journal of Law and Economics* 52 (1): 145–170.

Cason, Timothy N, Tatsuyoshi Saijo, and Takehiko Yamato. 2002. "Voluntary participation and spite in public good provision experiments: an international comparison." *Experimental Economics* 5 (2): 133–153.

Croson, Rachel, and Nancy Buchan. 1999. "Gender and culture: International experimental evidence from trust games." *The American Economic Review* 89 (2): 386–391.

Felice, Emanuele. 2013. *Perché il Sud è rimasto indietro.* Il Mulino.

Gibson, John, David McKenzie, Halahingano Rohorua, and Steven Stillman. 2014. "Beliefs, Preferences and Migration: Evidence from Combining Lab-in-Field and Natural Experiments." *Working paper.*

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2004. "The role of social capital in financial development." *The American Economic Review* 94 (3): 526–556.

Helliwell, John F., and Robert D. Putnam. 1995. "Economic Growth and Social Capital in Italy." *Eastern Economic Journal* 21 (3): 295–307. ISSN: 00945056, 19394632. http://www.jstor.org/stable/40325643.

Hong, Lu, and Scott E Page. 2001. "Problem solving by heterogeneous agents." *Journal of Economic Theory* 97 (1): 123–163.

Hoyman, Michele, Jamie McCall, Laurie Paarlberg, and John Brennan. 2016. "Considering the Role of Social Capital for Economic Development Outcomes in US Counties." *Economic Development Quarterly.*

Ichino, A., and G. Maggi. 2000. "Work environment and individual background: explaining regional shirking differentials in a large Italian firm." *The Quarterly Journal of Economics:* 1057–1090.

Knack, Stephen, and Philip Keefer. 1997. "Does social capital have an economic payoff? A cross-country investigation." *The Quarterly Journal of Economics:* 1251–1288.

Lazear, Edward P. 1999. "Globalisation and the market for team-mates." *The Economic Journal* 109 (454): 15–40.

Leonardi, Robert. 1995. "Regional Development in Italy: Social Capital and the Mezzogiorno." *Oxford Review of Economic Policy* 11 (2): 165–179. ISSN: 0266903X, 14602121. http://www.jstor.org/stable/23606534.

Nowell, Clifford, and Sarah Tinkler. 1994. "The influence of gender on the provision of a public good." *Journal of Economic Behavior & Organization* 25 (1): 25–36.

Ockenfels, Axel, and Joachim Weimann. 1999. "Types and patterns: an experimental East-West-German comparison of cooperation and solidarity." *Journal of Public Economics* 71 (2): 275–287.

Oosterbeek, Hessel, Randolph Sloof, and Gijs Van De Kuilen. 2004. "Cultural differences in ultimatum game experiments: Evidence from a meta-analysis." *Experimental Economics* 7 (2): 171–188.

Pronzato, Chiara. 2012. "An examination of paternal and maternal intergenerational transmission of schooling." *Journal of Population Economics* 25 (2): 591–608.

Putnam, Robert D, Robert Leonardi, and Raffaella Y Nanetti. 1993. *Making democracy work: Civic traditions in modern Italy.* Princeton University Press.

A Algorithm for the creation of groups

The following algorithm was implemented to subdivide participants of each session in four groups. Importantly, in each session, each school was represented by a maximum of 3 students.

1. Create three empty lists: $\mathcal{S}$(outh) with 6 slots, $\mathcal{N}$(orth) with 6 slots, $\mathcal{M}$(ixed) with 12 slots. A slot is occupied whenever a student is appended to a list.
2. If the session has strictly less participants from the North (South), remove one slot to the $\mathcal{N}$ ($\mathcal{S}$) list, respectively.
3. Let I be the school with the most students among schools still not processed.
4. Let $\mathcal{L}$ be the list $\mathcal{S}$ if the school is from the South, $\mathcal{N}$ otherwise.
5. If $\mathcal{L}$ has a free slot, append a randomly selected student from I to it.
6. If there are still students to be placed from I, append them to $\mathcal{M}$.
7. If there are still schools to be processed, go back to point 3.
8. Create two lists $\mathcal{M}1$ and $\mathcal{M}2$ from elements of $\mathcal{M}$ in odd and even positions, respectively.

The rationale for ordering schools by size was to guarantee that no two students from the same school would end up in the same group (i.e. that schools with more students, and hence more difficult to place, would "choose" - i.e. randomly assign their students - first).

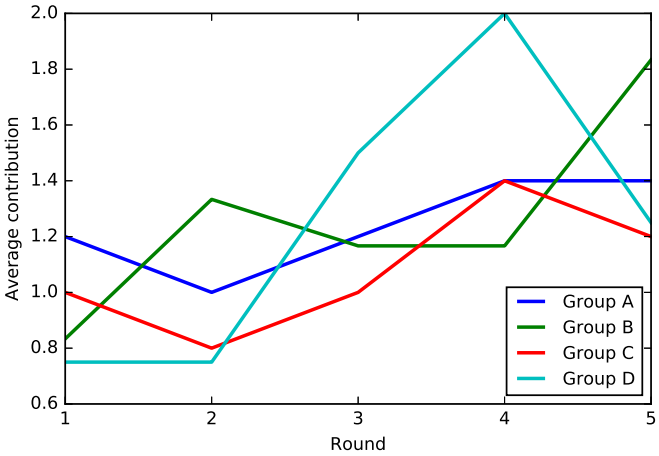
B Additional material

Figure 3 features an example of how information about past group contributions was shown to participants (from round 3 onwards).

Table 2 provides some descriptive statistics: for each session, we show the distribution of individual characteristics (geographic origin/gender) based on the assignment of individuals to the treatment. T-tests ran on the each session fail to reject the null of identical distribution between the two categories, with the exception of Session 2 ($p = 0.001$), in which homogeneous groups were composed only of female participants (we take this into account in Section 4.4).

Table 3 provides information about the 12 prize winners (three for each session). In theory, individual earnings could go from 2 to 32 (see Equation (2)). The signs of deviations between the shares of winners and the shares of sample presenting each feature are in line with results presented in the main text (females contribute more, although not significantly, "North-only" groups perform better, although not significantly, homogeneous groups perform better).

Figure 3: Example of past information, as shown to participants



Note: Information shown to participants of session 1 before the last round (labels translated from Italian).

Table 2: Descriptive characteristics

		Female		**North**	
Session	**Treatment**	**0**	**1**	**0**	**1**
1	0	5	4	6	3
	1	3	8	6	5
2	0	6	3	4	5
	1	0	10	4	6
3	0	5	4	6	3
	1	3	8	6	5
4	0	4	5	5	4
	1	3	7	5	5

Table 3: Descriptive characteristics of winners

Session	**Rank**	**Female**	**North**	**Treatment**	**Total gain**
1	1	0	0	1	22
1	2	1	1	1	21.4
1	3	0	1	0	21
2	1	1	0	1	24
2	2	1	1	1	23
2	3	0	1	0	22.5
3	1	0	0	0	23.6
3	2	1	0	1	22.7
3	3	1	0	1	22.7
4	1	1	1	1	22.8
4	2	0	1	1	22.8
4	3	1	0	0	22.5
Winners (share):		**7 (58%)**	**6 (50%)**	**8 (67%)**	
Share of all participants:		63%	46%	54%	

C Supplementary results

In the following, we combine Equations 2 and 3, interacting phase and treatment dummies with the geographic origin of participants.

$$x_{i,t} = \zeta^f F_i + \sum_{P=I}^{IV} T_{t,P}(\zeta_P + \zeta_P^h HOM_i + N_{i,P}(\zeta_P^n + \zeta_P^{hn} HOM_i)) + \epsilon_{i,t}. \quad (5)$$

Hypotheses (HdII), (HdIII) and (HdIV) allowed us to investigate whether being in a homogeneous group (instead of a heterogeneous one) has an effect on contributions. The estimation of Equation (5) can help us verify if there is a treatment effect *conditional on the geographic origin of individuals*. Namely, we can answer such question by running the following joint tests on coefficients presented in Table 4:

- $\mathcal{H}_0 : \zeta_P^h + \zeta_P^{hn} > 0$ for individuals from the North,
- $\mathcal{H}_0 : \zeta_P^h > 0$ for individuals from the South,

for each phase $P = II, III, IV$. From such tests, no significant differences emerge ($p = 0.183, 0.239, 0.259$ for the North, $0.149, 0.517, 0.204$ for the South, respectively).

By exploiting the disaggregation along the dimension of geography, we can also compare North-only and South-only groups between them. This is done by testing $\mathcal{H}_0 : \zeta_P^n + \zeta_P^{hn} > 0$ for each phase $P = II, III, IV$.[14] Results do not suggest that people from the North act differently from people from the South in homogeneous groups ($p = 0.703, 0.191, 0.306$, respectively).

By running the same analysis for mixed groups, we can instead compare the behavior of Southern and Northern individuals subject to the *same* treatment (i.e. being in a mixed group) in each phase.[15] Namely, we test $\mathcal{H}_0 : \zeta_P^n > 0$ for each phase $P = II, III, IV$: in line with Result 3, in mixed groups we find a higher level of contributions on behalf of Northeners compared to Southeners, for two phases out of three ($p = 0.669, 0.055, 0.087$, respectively).

14. It is worth stressing the fact that such tests pool together an intrinsic feature (the geographic origin) and a possible treatment effect (being in a North-only or South-only group).

15. We had already tested Hypothesis (Hn) on such a subsample, but pooling together all phases, that is not looking for an effect of design changes.

Table 4: Additional estimation results

		Eq. (4) (δ)			Eq. (5) (ζ)
				Female	0.104
					(0.076)
Phase	I	0.950***	Phase	I	0.926***
		(0.179)			(0.108)
	II	1.250***		II	1.046***
		(0.195)			(0.195)
	III	1.350***		III	1.236***
		(0.198)			(0.152)
	IV	1.225***		IV	1.212***
		(0.158)			(0.155)
Phase interacted with treatment	h,I	0.272	Phase interacted with treatment	h,I	0.032
		(0.263)			(0.156)
	h,II	0.194		h,II	0.341
		(0.277)			(0.224)
	h,III	-0.128		h,III	-0.135
		(0.312)			(0.203)
	h,IV	0.331		h,IV	0.270
		(0.232)			(0.203)
Phase interacted with Female	f,I	0.144	Phase interacted with North	n,I	0.098
		(0.229)			(0.088)
	f,II	-0.250		n,II	0.113
		(0.227)			(0.258)
	f,III	0.150		n,III	0.322*
		(0.293)			(0.155)
	f,IV	0.306**		n,IV	0.246*
		(0.115)			(0.134)
Phase interacted with F. and t	hf,I	-0.290	Phase interacted with N. and t.	hn,I	0.036
		(0.307)			(0.172))
	hf,II	0.321		hn,II	-0.050
		(0.308)			(0.309)
	hf,III	-0.039		hn,III	-0.069
		(0.361)			(0.244)
	hf,IV	-0.195		hn,IV	-0.088
		(0.202)			(0.201)
	N	468		N	468

Note: The dependent variable in each model is $x_{i,t}$. Row labels indicate coefficients subscripts: see the description of each block for the interpretation of the coefficients. Group-level clustered standard errors in parentheses.

***p< 0.01, **p< 0.05, *p< 0.10.

In conclusion, while we confirm the higher level of contributions of Northeners (Result 3) in mixed groups, we find no evidence that the treatment effect is related to the geographic origin of subjects. That is, we cannot explain Results 1 and 2 as the consequence of an interaction between the treatment and the geographic origin. However, we cannot exclude the possibility that such "non-result" is due to the low numerosity of observations in each of the subsamples considered.