Immigration, Integration and Cooperation: Experimental Evidence from a Public Goods Game in Italy

Johanna Gereke[a], Max Schaub[b] and Delia Baldassarri [c,d]

[a] Mannheim Centre for European Social Research (MZES), A5, 6 68159 Mannheim, Mannheim, Germany; [b]Berlin Social Science Center (WZB), Berlin, Germany; [c] New York University (NYU), New York, USA; [d] Bocconi University, Milan, Italy

ARTICLE HISTORY

Compiled May 10, 2021

ABSTRACT

Immigration has rapidly changed the demographic profile of most European societies, increasing ethnic diversity. A large cross-disciplinary literature comparing homogeneous and diverse communities provides mainly observational evidence for a link between high levels of ethnoracial diversity and lower levels of public goods provision. These results could be driven either by composition effects - immigrants having different baseline levels of cooperativeness, and thereby lowering average cooperation rates - or context/interaction effects, whereby individuals lower their cooperativeness in response to the presence of non-coethnics. To disentangle these effects, we conducted a lab-in-the-field experiment with a sample of Italians and immigrants from Morocco and the Philippines residing in Milan. In our public goods experiment, participants were randomly assigned to either homogeneous or ethnically mixed groups. We find that Italians behave similarly in both homogeneous and heterogeneous groups, thus contradicting arguments about the negative effects of diversity on the native population, while there are both compositional and interactional effects when considering the behavior of Moroccan and Filipino immigrants, respectively. Further, differences largely disappear when we consider only the behavior of more integrated immigrants. Taken together, our results highlight the need for a more nuanced understanding of cooperation in multi-ethnic communities.

CONTACT Johanna Gereke. Email: johanna.gereke@mzes.uni-mannheim.de

KEYWORDS

ethnic diversity, immigration, cooperation, public goods game, experiment

1. Introduction

Increasing immigration and ethnic diversity in European societies have raised concerns about the negative consequences of immigration on social cohesion and the provision of public goods such as education and health care (Stichnoth & Van der Straeten, 2013). Meanwhile, there is a wealth of studies across the social sciences investigating the relationship between ethnoracial diversity and social outcomes, including trust and public goods provision (Alesina, Baqir, & Easterly, 1999; Alesina & Ferrara, 2005; Alesina & La Ferrara, 2002; Habyarimana, Humphreys, Posner, & Weinstein, 2007; Miguel & Gugerty, 2005; Putnam, 2007). Several of these observational studies find a statistically significant but modest negative relationship of ethnic diversity on these outcomes.[1] However, questions remain concerning the possible spuriousness of this relationship, as well as the specific mechanisms that might explain the link between increased ethnoracial diversity and lower public goods provision.

The few explanations that have been advanced can be summarized in two broad categories: *contextual/interactional* effects and *compositional* effects. The most common interpretation of the relationship in terms of contextual/interactional effects argues that ethnoracial diversity has negative effects on public outcomes due to in-group favoritism or out-group hostility (Alesina & La Ferrara, 2002; Brewer, 1999; Putnam, 2007). In other words, a co-ethnic bias leads individuals to adjust their behavior depending on the identity of their interaction partner by favoring those with whom they share an ethnic identity because they attach higher utility to the well-being of their in-group than any out-group members (Alesina, Glaeser, & Glaeser, 2004). In the context of ethnic diversity due to immigration, contextual/interactional explanations often argue that the physical presence of immigrants in a community may make members of the majority population less pro-social, either towards immigrants specifically or towards everyone (Dinesen & Sønderskov, 2015; Enos, 2017; Gereke, Schaub, &

[1] For a recent meta-analysis on the link between ethnic diversity and social trust, see Dinesen, Schaeffer, and Sønderskov (2020).

Baldassarri, 2018; Laurence & Bentley, 2016; Legewie & Schaeffer, 2016). A second stream of research, based on a rigorous experimental design and employing behavioral games, has moved beyond in-group favoritism to investigate a broader set of mechanisms that might account for the ability of homogeneous communities to generate higher levels of public goods. For instance, homogeneous groups might be better at communicating, and thus coordinating, because of their common language, and shared social norms. Moreover, social sanctioning might be easier in homogeneous contexts, because of denser social networks. Finally, homogeneous groups are more likely to show similarity of preferences for specific public goods (Habyarimana et al., 2007).

In contrast to these contextual explanations, other explanations argue that the apparent association between ethnoracial diversity and lower public goods provision is simply a compositional effect (Abascal & Baldassarri, 2015; Bertocchi, 2016). Such compositional effects may appear because ethnicity often overlaps with socioeconomic characteristics of disadvantaged groups, such as poverty or lower levels of education (Abascal & Baldassarri, 2015; Sturgis, Brunton-Smith, Read, & Allum, 2011). For example, in the Brazilian context, Kustov and Pardelli (2018) find that diversity is not detrimental to public goods provision *per se* but only insofar as it reflects a larger share of the disadvantaged group in the local population, suggesting that compositional characteristics related to wealth and interpersonal trust might explain the racial demography effect in this case. Research has also shown that not only a person's own socioeconomic status is related to cooperation (i.e. higher SES is associated with higher contributions) but that expectations based on the SES of the interaction partner(s) also influence the willingness to cooperate (Bechtel & Scheve, 2017). For example, Zhang, Aidenberger, Rauhut, and Winter (2019) find lower helping rates in native-immigrant encounters, especially in cases involving stereotypically low-status immigrants in Switzerland. Similarly, Schaub, Gereke, and Baldassarri (2020) document in an online experiment with participants from the U.S., in which both the income (low vs. high) and racial identity (white vs. black) of the interaction partner varies, that cooperation is lower when people are paired with low-income partners, thus showing the importance of others' socio-economic status on expectations of cooperation. Besides these compositional diversity effects, others may also stem from

different life experiences, whether these are related to immigrants' treatment in the host country or to differences in the level of public goods provision in immigrants' origin countries.

Differentiating between contextual and compositional effects is important because each leads to different implications for how to best address the impact of ethnoracial diversity in contemporary multicultural societies. While the first perspective focuses on intergroup dynamics, the second perspective focuses on the structural and historical conditions that may have led different groups to exhibit different baseline levels of cooperativeness. We therefore set out to test the following research question: Is cooperation in contributing towards a public good lower in ethnically mixed groups than homogeneous ones? And, if so, is it because of individuals' cross-cultural or socioeconomic differences (compositional effect) or because of their reaction to out-group members (contextual/interactional effect)?

A major difficulty in comparing homogeneous and heterogeneous communities concerns the fact that in Western European and North American countries, we are usually only able to observe either homogeneous native communities or ethnically mixed ones. Thus, observational research mostly measures the behavior of natives in homogeneous and mixed communities and the behavior of immigrants, or more generally ethnic minorities, in mixed groups. Missing information on the behavior of minorities in homogeneous groups makes it impossible to disentangle compositional from contextual effects. Namely, as Abascal et al. puts it in the U.S. context, "[w]ithout homogeneously Non-White communities, it is impossible to determine if homogeneous communities (and their residents) are better off because they are homogeneous, or because they are predominately White (and, more generally, advantaged)" (Abascal, Ganter, & Baldassarri, 2021, p. 4) . Further, the measurement of the behavior of natives in homogeneous and ethnoracially mixed groups in the real world is often vulnerable to self-sorting into neighborhoods, associational groups and other contexts of immigrant-native interactions depending on previously existing differences in preferences for contact with ethnic minorities (Crowder, Hall, & Tolnay, 2011).

To overcome these limitations of observational research, we run a controlled large-scale lab-in-the-field experiment in Italy. To disentangle compositional from interac-

tional effects in public goods provision we included both Italians and immigrants from two ethnic minority groups, Filipinos and Moroccans. Using a standard public goods game, we control the ethnic/national composition of the group and randomize participants into either ethnically mixed or homogeneous groups, including homogeneous ethnic-minority groups. We find that Italian natives behave similarly in both homogeneous and heterogeneous groups, thus contradicting arguments about the negative effects of diversity on the native population. Cooperation rates are slightly lower in mixed groups than homogeneously Italian groups due to immigrants' behavior. Moroccans contribute less than other ethnic groups regardless of whether they are in homogeneous or heterogeneous groups (compositional effect), while Filipinos have higher levels of cooperation in homogeneous settings than in mixed Italian-Filipino groups (interactional effect). Taken together, these findings support the conjecture that the effect of diversity differs depending on the specific ethnic groups in question and their ability to produce a public good. Furthermore, we also find that the diversity gap narrows and disappears when we consider only the behavior of more integrated immigrants, suggesting that economic, political and social integration may be a pathway to higher public goods provision in ethnically diverse societies.

This article breaks new ground in the research on the effects of ethnic diversification on public goods provision by providing evidence concerning the behavior of both natives and immigrants in homogeneous vs ethnically diverse groups. The ethnic diversity literature has so far focused too much on a narrow, and quite misleading, definition of diversity as the share of natives in a community, and on native responses to outgroups. We argue that more attention should be paid to how minorities themselves behave and to important differences in the ability to produce public goods between minority groups. A main contribution of this article is the broader finding that differences among immigrant groups exist and can have important consequences for the effect of ethnic diversity on a community's collective action potential. Moreover, our work connects with a larger literature on intergroup relations in diverse societies arguing that we need to pay more attention to the level of social and economic integration of immigrants and ethnic minority groups in modern multiethnic societies to assess their impact on public goods provision (Baldassarri & Abascal, 2020, cf.).

2. Ethnoracial Diversity and Public Goods Provision

Most of the observational literature on the effects of ethnic diversity on public goods provision embraces the intergroup perspective in social psychology, which examines the dynamics of contact between people from different groups (Allport, Clark, & Pettigrew, 1954; Pettigrew & Tropp, 2013). This body of research on intergroup contact has documented that an in-group bias is a persistent feature of human behavior (Tajfel, Turner, Austin, & Worchel, 1979). Evolutionary theory has suggested that the coevolution of in-group favoritism and out-group hostility emerged during violent intergroup conflict (J.-K. Choi & Bowles, 2007). Nowadays, negative reactions of the majority native majority population towards immigrants (out-groups) is often explained with this in-group/out-group paradigm (Enos & Gidron, 2018). However, much of this scholarship is so far "group-blind", meaning different ethnic groups are treated similarly as indistinguishable out-groups. Meanwhile, some scholars have highlighted that ethnic groups are rarely interchangeable but rather differ greatly in terms of their migration histories, economic interdependence and market integration, and/or position in the social and ethnic hierarchy of a society (for a similar critique see, Baldassarri and Abascal (2020); Kustov and Pardelli (2018); Sidanius and Pratto (2001)).

However, some scholars have cautioned that the negative association may simply be an artefact reflecting the different composition of homogeneous and ethnically mixed communities in many countries. For example, homogeneous and mixed communities often differ in terms of important socio-economic characteristics such as concentrated poverty, educational attainment and residential instability (Abascal & Baldassarri, 2015; Sturgis et al., 2011). There is some evidence from U.S. communities that suggests that economic factors are actually stronger predictors of collective outcomes than ethnoracial indicators (Portes & Vickstrom, 2011; Sampson, 2012; Schaub et al., 2020). Furthermore, as Abascal and Baldassarri (2015) have shown in the U.S. context, once taking into account the lower share of whites in mixed communities and their higher scores on prosocial indicators such as trust, negative associations with ethnic diversity disappear. Similarly, Kustov and Pardelli (2018) find that homogeneous Afro-descendent communities in Brazil have lower public goods provision than

mixed communities even after taking into account the disproportionate poverty, environmental conditions and underdevelopment that mark many of these communities. Their findings thus also call into question the often assumed universal benefit of ethnic homogeneity for public outcomes.

Empirically, it is difficult to tease apart the effects of ethnoracial diversity on public goods provision because of several limitations that characterize most previously studied contexts. First, as Kustov and Pardelli (2018, p.1) have also argued, it is crucial to be able to compare outcomes for diverse and homogeneous communities of all ethnic groups. Instead, most of the existing literature is based on comparisons of mixed communities to just a single homogeneous group, usually the dominant group or majority population in a society. However, the lack of comparison of diverse communities to homogeneous ethnoracial minority groups does prevent previous research from examining to what extent potential lower contributions of ethnic minorities may be explained by compositional or contextual factors. Second, another problem that observational research on ethnic diversity faces is the possibility that (particularly) the majority population self-selects into diverse or homogeneous communities based on unobserved preferences (Crowder et al., 2011). As a result of this self-sorting, any association found between ethnic diversity and public goods provision may be confounded by this selection issue. To overcome these limitations, we conduct a controlled lab-in-the field experiment that allows us to randomize participants into homogeneous and mixed groups (thus eliminating the possibility of self-selection) and also provides us with the means to compare the behavior of all ethnic groups between mixed and ethnoracially homogeneous sessions.

3. The Italian Context

The motivation for examining the effect of ethnoracial diversity in Italy is twofold. First, although Italy has only relatively recently become a country of immigration, it is now one of the major destinations for international migration in Europe (Ambrosini, 2013). Foreign-born residents comprise approximately 10 per cent of the Italian population (OECD, 2019). Among non-EU foreigners holding a residence permit in

Italy, the largest groups were Moroccans (454,817) followed by Albanians, Chinese, Ukrainians and Filipinos (162,469) as of 2017 (Istat, 2017).

Second, Italy's Mediterranean shores have also become a common entry point of African and Asian immigrants into the European Union, and thus the issue of immigration is highly politicized in the Italian political system influencing political debates about concerns over public spending and redistribution (Alesina, Miano, & Stantcheva, 2018). At the same time, economic demands for cheap foreign labour in low-skilled sectors, such as agriculture, construction and elderly care, have persisted even during recent financial recessions (Venturini & Villosio, 2018).

Anti-immigrant stereotypes are widespread in Italy (Alesina et al., 2018), leading some early immigration scholars to argue that the country is a unique case within Europe in terms of the high levels of xenophobic attitudes among the population and an early rise of explicitly anti-foreign political parties (Sniderman, Peri, de Figueiredo Jr, & Piazza, 2002). However, Italians today seem to differentiate between different outgroups: immigrants and minorities from Africa are, at least in certain settings, confronted with more negative attitudes and behavior than Asian immigrants (Zhang, Gereke, & Baldassarri, 2020).

While immigrants in Italy are usually employed in low-wage jobs, there is considerable heterogeneity among the occupational sectors in which the different immigrant groups are represented. In our study, participants were native (white) Italians, Filipinos, and Moroccans residing in Milan. Filipino immigrants in Italy are mainly employed in domestic care positions (e.g. elderly care). In Italian society, female immigrants in particular "complement/substitute the caring activities of women in the family and insufficient public services since the 1990s" (Del Boca & Venturini, 2016, p.19). As such, many immigrant women from the Philippines have become what scholars have labeled "intimate foreigners" with "a deep-seated familiarity with the daily lives of their Italian employers but [who] yet remain outsider in Italian society" (Parreñas, 2016, p.110). Moroccans on the other hand are seldom employed in private households and more often work in the construction or service sectors. This difference is also reflected in our own post-experimental survey results, in which 71% of the Filipino participants compared to only 39% of our Moroccan participants report being

employed by an Italian. As a result of this occupational segregation, it is reasonable to assume that Moroccans tend to have less personal contact (at least in their workplaces) with Italians than Filipino immigrants.

Both Moroccan and Filipino immigrant communities are well-established in Italy, exhibiting similar patterns when it comes to the duration of their stay. In fact, the Filipino community has the longest mean duration of stay among all immigrant groups in Italy (more than 12 years), followed by immigrants from Morocco with a mean duration of 10 years (DAgostino, Regoli, Cornelio, & Berti, 2016). However, this similarity masks important differences in various dimensions, such as the share of female immigrants: among the Filipino community 41% are women, while among the Moroccan community the share of women is only 14%. Also, the level of education among the two immigrant groups differs significantly: Moroccans have a much higher share of individuals who have primary or lower secondary education only (70% vs. 44% among Filipinos) (DAgostino et al., 2016, p.88). Lastly, the majority of Moroccans are Muslim, while most of the Filipinos share a common Roman Catholic faith with Italians. This difference may be relevant to comparing inter-group behavior because prior research has highlighted the role of religion, and particularly Muslim faith, as defining factor for ethnic boundaries in today's European societies (Adida, Laitin, & Valfort, 2016a).

One additional but important reason for focusing our study on these two different immigrant groups in particular is that most immigrants from these two countries are visually distinct from the average Italian. This difference in physical appearance is crucial in our design as it is our main tool to signal the group composition (ethnically homogeneous or mixed) to other participants in the multi-player public goods game.

4. Experimental Design and Implementation

4.1. *Experimental evidence on ethnic bias and intergroup cooperation*

Besides field experimental methods, such as correspondence and audit tests that have been increasingly employed to assess the level of discrimination against immigrants and ethnic minorities in employment or housing markets (Auspurg, Schneck, & Hinz, 2019)

or in everyday situations (D. D. Choi, Poertner, & Sambanis, 2019; Winter & Zhang, 2018; Zhang et al., 2019, 2020), there is a growing body of experimental research across the social sciences that examines ethnic bias and intergroup behavior using economic experiments largely developed to capture cooperation dynamics (Adida et al., 2016a; Bader & Keuschnigg, 2020; Berge et al., 2020; Blum, Hazlett, & Posner, 2020; Cettolin & Suetens, 2019; Fershtman & Gneezy, 2001; Gereke, Schaub, & Baldassarri, 2020; Habyarimana et al., 2007; Marshall & Paler, 2020; Schaub, 2017; Schaub et al., 2020). The increasing use of monetary incentivized behavioral games as a measurement tool for ethnic bias has also come in response to the growing awareness and concerns that traditional survey research on this topic is largely unreliable because of respondents not truthfully self-reporting their prejudices due to social desirability concerns (for a similar critique, see for example (Enos & Gidron, 2016)).

Most political scientists have so far largely focused on studying ethnic bias in public goods provision using behavioral games in the Sub-Saharan African context, where ethnicity is particularly salient in political and social interactions (Berge et al., 2020; Habyarimana et al., 2007; Miguel & Gugerty, 2005). Related work in India has examined public goods contributions in religiously homogeneous and heterogeneous groups using the Hindu-Muslim cleavage and found that contributions across the standard public goods game did not systematically differ across group composition (Keuschnigg & Schikora, 2014; Tusicisny, 2017). There is also some recent work in Lebanon, where researchers used a six person public goods game similar to our experimental set-up, to study the effect of class and sectarian cleavages on women's collective action capacity by randomly assigning participants to same-gender groups that were homogeneous and heterogeneous in terms of class and sectarian compositions (Marshall & Paler, 2020). In the European context, where ethnoracial diversity is increasing (at least in urban areas) due to globalization and immigration dynamics, researchers have so far failed to examine what happens to cooperation when communities become increasingly ethnically diverse due to immigration in a controlled study environment. We therefore hope to begin filling this research gap with the current experimental study using a six person public goods game that was conducted in the multi-ethnic metropolis of Milan.

	Italians	Filipino	Moroccans
N Subjects	561	140	193
Female	54%	61%	26%
Median Age	41	37	39
Employed	70%	62%	23%
No high-school degree	4%	23%	51%

Table 1. Demographic Differences of Study Participants by Nationality (see Table 3 for Summary Statistics of Full Control Variables)

4.2. *Data and Methods*

Our data was collected through a large-scale lab-in-the-field experiment in Milan, Italy, with a sample of 561 Italian, 140 Filipino and 193 Moroccan participants.[2] In contrast to standard laboratory experiments in which identities are often created ex novo, we test the emergence of cooperation with preexisting ethnic/national groups. Previous research using behavioral games suggests that group identity, shared norms and past experiences are important for decisions that include the weighing of private and public interests as in our set-up (Baldassarri, 2009; Henrich et al., 2004).

We recruited a representative sample of Italians (broadly representative of the population of Italy in terms of basic socio-demographic characteristics) and immigrants from Morocco and the Philippines residing in Milan who participated in 149 sessions of six participants each (see see Table 1). For the recruitment of the sample, we considered three neighborhood criteria: ethnic diversity, wealth, and urbanity. Further, in selecting neighborhoods to sample from, we excluded those with extreme values on any of these three dimensions. In total, our neighborhood sample is made up of 55 out of 88 neighborhoods of Milan (NIL) with 5-28% foreigners, not extreme poverty, unemployment or wealth, and urban rather than suburban structure in terms of population density. [3] All the sessions took place in a central and easily accessible location in Milan at the offices of a well-known survey and marketing firm during the fall of 2017.

Participation took around 90 minutes and was compensated with a minimum of 15 and a maximum of 25 euros paid in cash, depending on individuals' behavior in the

[2]Prior to data collection, the study received ethical approval from the ethics committee of the grant hosting institution (Bocconi University).

[3]More information on our sampling strategy and recruitment can be found in Section 7.2 in the Appendix.

behavioral games.[4] On average, participants received 20 euros in cash in addition to a guaranteed voucher worth 10 euros paid to each participant as a show-up fee.[5]

4.3. *The Public Goods Game*

Our outcome of interest is cooperative behavior in a Public Goods Game. The Public Goods Game (PGG) is a standard behavioral measure to capture individuals' behavior in real-world collective action dilemmas. In contrast to other behavioral games used to commonly measure in-group vs. out-group solidarity and trust, such as the dictator or trust game, the PGG captures the willingness to cooperate in the production of public goods in a context in which individuals are faced with a trade off between their own individual gain and the collective wellbeing. Public goods, being notoriously non-excludable and non-rivalrous,[6] give rise to the free-rider problem: participants have little rational incentive to contribute to the production of the collective good, as it is always more convenient for them to free ride and let other participants contribute to the public good provision. However, if all participants follow this pattern, no public good is ever produced. The payoff structure of the PGG captures this dynamic.

Participants in the same session were randomly assigned to either homogeneous groups or to ethnically mixed groups consisting of either four Italians and two Filipino participants or one to three Moroccans and three to five Italians, so that there were always six participants in a room (see Table 2 for an overview of group compositions and Table 4 in the Appendix for balance checks of our randomization). For the mixed sessions, in total there were 34 sessions with Filipinos and 9 sessions with one Moroccan and five Italians, 35 sessions with two Moroccans and four Italians and 14 sessions with three Moroccans and three Italians.[7]

A key feature of our design is that we purposely provided no explicit cues about the nationality of the other participants. Instead, the composition of the group was only

[4] At the beginning of each session, all participants provided written informed consent.

[5] Participants final cash payment was determined by one randomly selected individual task and one randomly selected round of the PGG. The average hourly wage in Italy was 21 euros in 2017. While there is no official minimum wage in Italy, most of the minimum wages determined by unions are 7 euros per hour. We therefore believe that we provided fair compensation for participants' time and effort.

[6] Non-excludable means that everyone benefits from the public good regardless of their own contribution. Non-rivalrous implies that the usage by one individual does not reduce its availability to others in the group.

[7] We also tried to ensure that out of the six participants, there was a relatively balanced gender ration of 3-3 or 2-4.

Condition	Group Composition	Nationality	Session N	Participants N
1	Homogeneous	Italian	33	198
2	Mixed	Italian+Filipino	34	204
3	Mixed	Italian+Moroccan	58	348
4	Homogeneous	Filipino	12	72
5	Homogeneous	Moroccan	12	72

Table 2. Treatment Conditions in PGG Experiment

observable via the visual cues of seeing the other participants while entering and during the session. For this reason, we picked two ethnic minorities that are easily visually distinguishable from native Italians. In following this approach, we differ from other experimental studies where participants are provided with pictures, names, or other direct information about the ethnic background of their fellow participants (Adida, Laitin, & Valfort, 2016b; Gereke et al., 2020; Habyarimana et al., 2007).

Each session had a moderator who welcomed participants and orally and visually provided several examples before the start of the public goods game to ensure that participants understood the payoffs and the dynamics of the game. This introduction was followed by three control questions to measure participants' understanding of the rules of the game.[8]

We focus on the very first round of the PGG because participants do not have any peer feedback on which to orientate their expectation of others' contributions.[9] In essence, this represents most of our common, one-time encounters in modern societies and allows us to answer our research questions, namely (1) Is the cooperative behavior of natives and immigrants affected by whether they are in ethnoracially homogeneous vs heterogeneous communities? and (2) to what extent do compositional or contextual/interactional factors explain any potential differences?

In the PGG, players decide anonymously how to divide 10 EUR between a private and a public account. The contributions that are paid into the public account are doubled by the experimenters and then redistributed evenly across all participants.

[8]The majority of our participants, 72% successfully answered most control questions correctly. The control questions were: the maximum amount a player could contribute to the PGG, the minimum amount a player could contribute, and the total number of players in the game. However, there were some differences in the understanding of the game across ethnic groups as shown in detail in the Summary Statistics Table 3 in the Appendix.

[9]In total, the PGG was played for 8 rounds but participants only knew that they will be interacting with the same group for several rounds, not the exact number of rounds. All participants played two initial rounds of the standard public good game as described in the paper. Afterwards some sessions were introduced to different punishment treatments. The PGG was the first group activity in each session, which always started with a short survey on socio-demographic characteristics and several individual decision tasks.

Figure 1. Experimental Room Set-up

Participants were seated around a large conference table and made their decisions behind a three-sided cardboard screen that allowed for some privacy but also ensured that participants could see each other. The game was programmed using oTree (Chen, Schonger, & Wickens, 2016) and decisions were made on a tablet (see Figure 1).

As soon as everyone in the room entered their decision on the tablet, all contributions become common knowledge. Each participant was presented with a feedback screen of all contributions without being able to match contributions to an individual's identity in the room, as participants received a random number (1 through 6) which was unknown to the others (see Figure 2). They were also presented with their personal payoff, which represents the sum of their private account plus the returns from the public account. No communication was possible between the participants before and after entering the room.

4.4. *Main results*

First, we will present descriptive evidence of the variation in cooperation between homogeneous Italian and ethnically mixed groups before comparing all mixed groups by nationality and the average of all homogeneous groups and mixed groups in Figure 3. Next, we compare homogeneous and mixed groups by nationality separately in Figure

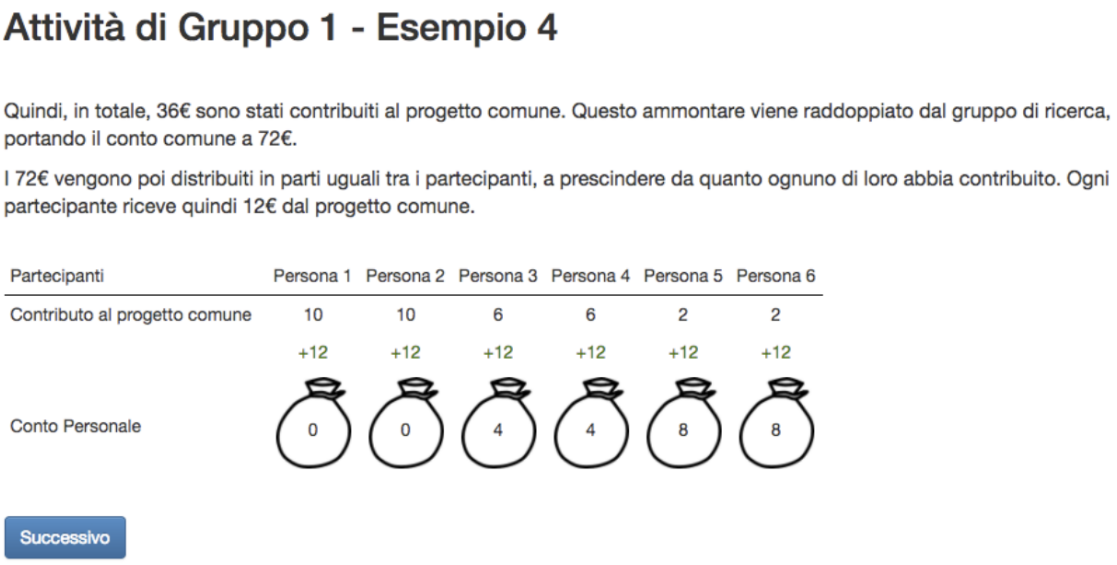


Figure 2. Example Screenshot of PGG feedback

4 before moving to additional analyses of contribution levels to the PGG by Italians and immigrants of different socio-economic and political integrational levels. To test for statistical significance, we estimate t-tests and ordinary least squares (OLS) regression models using robust standard errors.[10] We estimate all results with and without control variables to account for potential confounding. The control variables include basic socio-demographics, such as age, gender and level of education as well as game-specific and implementation variables, including participants' ability to understand the game instructions (based on three control questions) as well as whether they personally knew someone who had previously participated in the study.

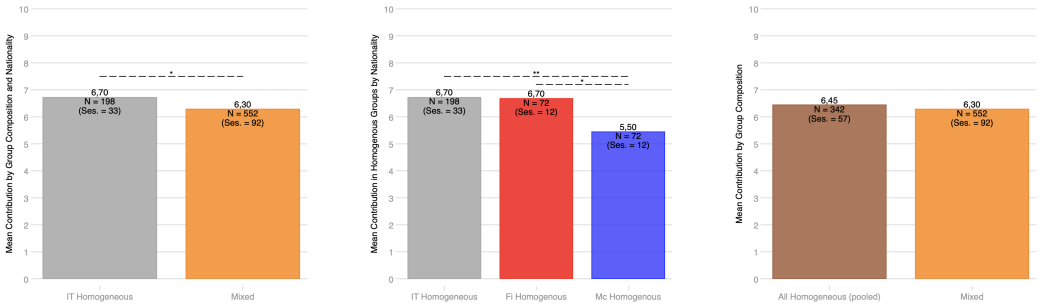
We start by presenting averages in contributions in the PGG by group composition. If we simply compare contributions in homogeneous Italian groups to contributions in mixed groups (Figure 3a) we find that mixed groups with on average four Italians and two immigrants produce overall contributions to a public pot of 6,30 EUR, while homogeneous Italian groups achieve contributions of 6,70 EUR, a difference which is marginally statistically significant at the 10% level (p-value = 0.106, see Table 5 in the Appendix). At a first glance, this result resonates with much of the observational

[10]We use robust standard errors because this is this is the first activity and round of the PGG, in which participants engaged together and no prior communication was allowed, so that we do not need to cluster standard errors at the group level. Nevertheless, our results are robust to different estimation strategies, such as clustering SES at the session level or estimating random intercepts in our models.

literature. However, a closer look at the data questions whether diversity is indeed at the origin of lower public goods provision. Specifically, the difference in contributions between homogeneous Italian and ethnically mixed groups disappears once we include control variables for social inequalities, such as the level of education (see Model 3 Table 5). This suggests that compositional effects seem to be driving most of the difference in behavior.

Second, our design, which includes homogeneous sessions with Filipino and Moroccan immigrants, allows us to determine whether group homogeneity in itself may lead to higher levels of public goods provision, or whether the observed small, but marginally statistically significant difference in contributions between ethnically mixed and homogeneous Italian groups should be better interpreted in light of Italians' socio-economic advantage. First, we compare contributions in homogeneous groups. As Figure 3b shows, homogeneous Italian and Filipino groups contribute at similarly high levels of 6,70 EUR but homogeneous Moroccan groups contribute on average significantly less (5,50 EUR). This difference in contribution levels between homogeneous Italian groups and homogeneous Moroccan groups is also statistically significant at the 5% level and robust to including socio-demographic and game-specific control variables (see Table 6 in the Appendix). Echoing previous findings from Brazil (Kustov & Pardelli, 2018), we conclude that homogeneity per se is no guarantee of high levels of solidarity and contributions to the public good. This finding becomes even more clear when we look at average contributions in all homogeneous groups and compare them to the mean contributions in all mixed groups as Figure 3c illustrates (see Table 7 in the Appendix).

Next, we investigate whether the lower contributions in the ethnically mixed groups are due to compositional or rather due to contextual/interactional effects. To do so, we compare contributions in homogeneous vs. mixed groups for each ethnic group separately to examine whether participants adjust their behavior to group composition. Figure 4 shows the differences in contributions to the public good by nationality and group composition. The results are shown separately for Italians (grey) in Figure 4a, for Filipino participants (red) in Figure 4b and for Moroccan participants (blue) in Figure 4c. The results indicate that Italians and Moroccans contribute similarly



(a) Homogeneous IT vs. Mixed Groups

(b) Homogeneous Groups by Nationality

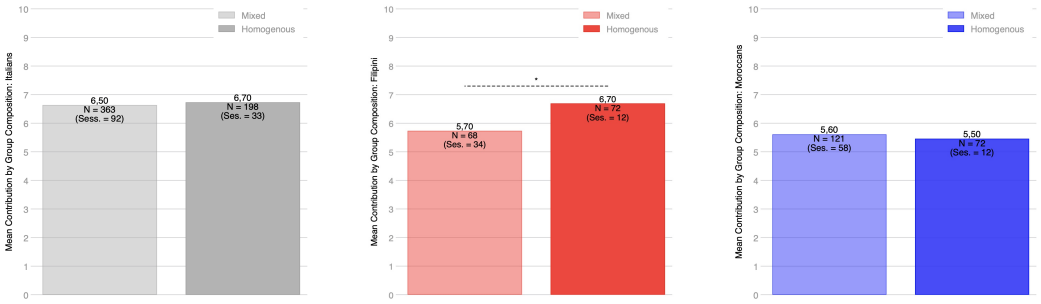
(c) All homogeneous vs. all Mixed Groups

Figure 3. Panel 3a shows average contribution levels in the public goods game by group composition for homogeneous Italian (grey) versus mixed (orange)) groups. Panel 3b shows average contribution levels for homogeneous groups by nationality: Italian (grey), Filipino (red), and Moroccan (blue). Panel 3c shows average contribution levels by group composition for all homogeneous (brown) versus mixed (orange)) groups. The results are drawn from Appendix Tables 5 6, and 7. Statistically significant differences between groups are indicated by **p $<$ 0.05, *p $<$ 0.10 above the dashed lines.

in homogeneous and mixed sessions. In other words, neither Italians nor Moroccans change their behavior when moving from homogeneous to heterogeneous groups, thus not showing much ingroup favoritism. We thus interpret the lower contributions in mixed groups compared to homogeneous Italian groups as mainly a compositional effect.[11] In contrast, Filipino participants make higher contributions in homogeneous (6,70 EUR) than in mixed Italian-Filipino groups (5,70 EUR) (but only marginally statistically significant, at the 10% level, see also Models 4-6 in Table 8). In our study, Filipino participants are the only ones that seem to adjust their contribution levels to the ethnic make-up of their PGG group, suggesting that there is a contextual or more specifically an interactional effect for this immigrant minority group.

We therefore find that diversity has little effect on the cooperative behavior of Italians, and depending on the immigrant minority, different mechanisms explain lower overall contributions in the mixed public goods groups compared to homogeneous Italian groups: an interactional effect explains lower contributions of Filipinos in mixed groups, while a compositional effect accounts for lower public goods contributions in Moroccan-Italian mixed groups.

[11]In the Appendix, Table 9, we examine whether Italians behave differently when matched with Filipinos or Moroccans. We find that Italians contribute less to the PGG when mixed with Moroccans than Filipinos. This difference is marginally significant (10% level). Lower cooperation rates seem to be driven by Italians in mixed groups matched with two or more Moroccan men. This is consistent with other research on the effect of the size of the outgroup (Adida et al., 2016b) and their gender (Gereke et al., 2020).



(a) Italians (b) Filipinos (c) Moroccans

Figure 4. Average contribution levels in the public goods game by mixed (transparent shading) vs. homogeneous (solid shading) composition for each of the three nationalities: (Italian (grey), Filipino (red), Moroccan(blue). Estimation results on which these figures are based are available in the Appendix, Table 8. A statistically significant difference between mixed vs. homogeneous groups for Filipinos is indicated by *p $<$ 0.10, above the dashed line.

4.5. *Do socio-economically integrated immigrants behave more like native Italians?*

Finally, we examine how the level of immigrants' incorporation in the whole society is related to their contribution to the public good. This is important from a public policy perspective because it allows us to cast some light on the trajectories of immigrants as they are experiencing differentiation in increasingly multi-ethnic European societies. Are these differences between the majority population and ethnic minorities likely to last or fade away in the future? Our hypothesis is that immigrants who are more socially, economically and politically incorporated into Italian society also have more similar contribution levels to Italians than immigrants who are not (yet) as well integrated, either because they have absorbed similar behavioral norms, are more similar to Italians in their socio-economic status, or they increasingly face similar structural conditions as Italian natives. It is important to note here that we did not randomize immigrants depending on their level of integration into homogeneous or mixed sessions although our random assignment of participants to different group compositions should have prevented any systematic differences. Therefore, the evidence presented here does not show causal effects but simply statistical associations.

We ask whether the small but statistically significant differences in cooperation rates are reduced if we compare contributions of Italians with those of immigrants who are already well integrated into Italian society. We address this question using a range of

socio-economic and political integration indicators, such as whether respondents are currently employed (as an indicator of economic integration), whether they hold Italian citizenship (political integration), whether they self-report speaking Italian fluently, and their level of social contacts with regards to the number of Italian friends and acquaintances (social integration).

As shown in Figure 5 and Table 10, we find that immigrants who are employed give significantly more than immigrants who are not participating in the labor market (P-value<0.05). Furthermore, we also see that the contributions of Italians who are employed and those who are not employed do not differ statistically from the contributions of immigrants who are employed. This sub-analysis seems to suggest that only immigrants who are currently not actively participating in the labor market make significantly lower contributions than any other group, hinting at the importance of employment and market integration for pro-social behavior (Baldassarri & Abascal, 2020).

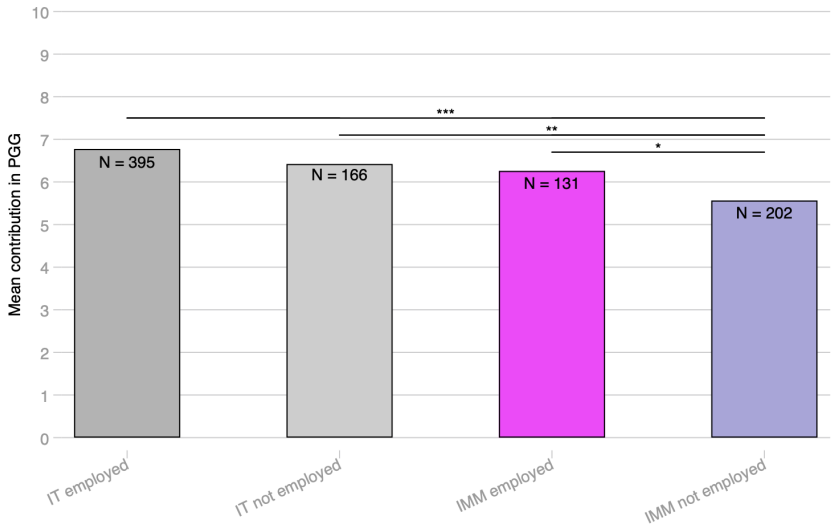


Figure 5. Average contributions in the public goods game, by nationality and employment status. Statistically significant differences are indicated by ***p < 0.001, **p < 0.05, *p < 0.10 above the lines. See Table 10 in the Appendix for full regression results.

In addition, Figure 6 shows that immigrants with Italian citizenship or those who speak Italian fluently do not differ from native Italians in their contribution levels to the public good. Similarly, immigrants who report having more contact with Italians in their social networks show more similar contribution levels to Italians than those with

fewer ties to the majority population, although for Italian acquaintances the confidence interval no longer overlaps with zero. This implies that they are statistically different from the contributions of Italians, though the difference is smaller for those with more than 10 Italian acquaintances compared to those with fewer Italian acquaintances. The OLS regression models reported in Table 11 in the Appendix comparing contributions of immigrants by their level of integration also confirm that immigrants who self-report greater levels of political and social integration contribute significantly more in the public goods game than immigrants who are less integrated.

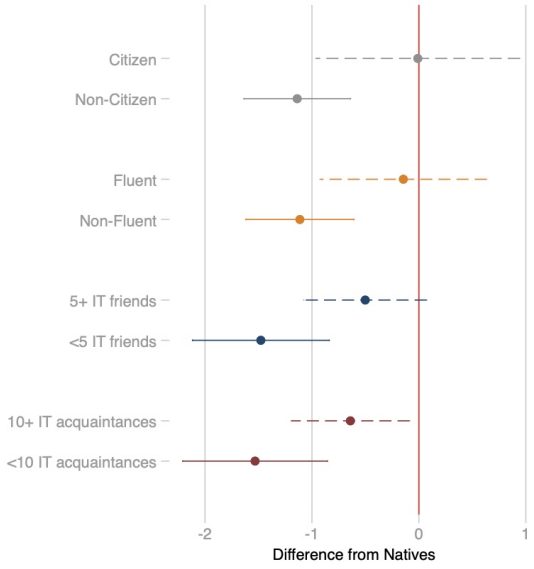


Figure 6. The figure displays estimates and 95% confidence intervals from ordinary least squares (OLS) regression results of the difference in contributions to the public good between Italians and immigrant participants by immigrants' integration level (N=894). For full regression results comparing immigrants by their level of integration, see Table 11 in the Appendix.

These results overall support our hypothesis that immigrants who are well integrated in Italian society show more similar contribution levels to the public good with Italians on average than with other immigrants who are less well integrated into Italian society. In sum, we find that higher levels of immigrants' social, structural and cultural integration into Italian society are associated with higher contributions to the public good.

5. Conclusion and Discussion

As societies become increasingly ethnically diverse, there is growing interest in understanding how ethnoracially heterogeneous communities can successfully cooperate and provide public goods. Previous studies have found an overall negative association of ethnic diversity on public goods provision (Alesina et al., 1999), but there is still a lack of evidence for a causal link and a debate about how to explain the oft-cited "diversity deficit" (i.e. the association between ethnic diversity and a variety of negative social outcomes).

We conducted a controlled public goods experiment with "native" Italians and immigrants from two different origin countries to test whether contextual and more specifically interactional effects based on adjusted expectations due to the ethnic/ national identity of the other group members can account for any "diversity deficit" in ethnically mixed communities. In showing that the majority group – Italians – do not cooperate less in ethnically mixed groups than in homogeneous groups, our results challenge the explanation that in-group favoritism can account for lower contributions in modern multi-ethnic communities. While this lack of discrimination among Italian participants in ethnically mixed groups may seem surprising given widespread anti-immigrant attitudes in Italy (Castelli Gattinara, 2017), related research in Italy using variation in foreign versus Italian names in a recent nationwide lost-letter experiment also does not find behavioral evidence of discrimination (Baldassarri, 2020). Furthermore, results from another study examining individuals interpersonal trust and other pairwise strategic intergroup interactions with the same sample as here suggest that ethnic discrimination does occur, just not in the interaction itself but instead when faced with a choice of whom to interact with (WITHHELD, 2021).

Similarly, we find no interactional effect for Moroccan participants in the public goods game. In other words, Moroccans also did not show any in-group bias, contributing similarly - but importantly in lower amounts than Italians - in homogeneous as in mixed groups. The lower level of contributions in homogeneous Moroccan groups speaks to the often overlooked fact that some homogeneous communities, especially those that are more socio-economically disadvantaged and have lower levels of educa-

tional attainment, can have poorer public good outcomes in ethnically homogeneous communities than in heterogeneous ones (see also Kustov and Pardelli (2018) for evidence in the Brazilian context). Overall, our results suggest that it is not ethnic diversity per se that undermines cooperation in multiethnic societies but rather that the conditions that immigrants often find themselves in lead to lower cooperation.

The findings also highlight that ethnic minority groups are not all equal in their cooperation behavior. In particular, we find that while neither Italians nor Moroccans differentiated between contributions in homogeneous and mixed groups, Filipino participants did display in-group favoritism.[12] Our results suggest that it might not always be the same mechanism (composition or contextual/interactional effects) that explains cooperation rates in multi-ethnic groups. Instead, depending on the makeup of the group, the history of a particular ethnic /national group and asymmetries in wealth, education and power between groups, we may find different cooperation dynamics. In this specific context, we conjecture that compared to Moroccans, Filipinos in Italy have more close-knit ethnic networks and community organizations (i.e. often linked to the church or migrant NGOs, such as Athika Overseas Workers and Communities Initiatives, Inc.) and that associational membership fosters in-group cooperation. At the same time, the Filipino participants in our study indicated that they were much less likely to plan to remain living in Milan for the rest of their life than the Moroccan participants (only 40% vs. 67% among Moroccans). This difference in permanent settlement and future migration plans may explain the lower contributions of Filipino participants in ethnically mixed groups as a result of lower levels of attachment to the local community. Future research should more closely study the mechanisms behind differences in lower public goods contributions among different minority populations.

While this study advances our understanding of the diversity effect on public goods provision, there also remain several limitations of the current design that should be addressed in future work. First, we did not test cooperation in multi-ethnic contexts with more than one immigrant minority present at a time. However, modern societies and communities are usually characterized by having immigrants from various countries simultaneously present and not only one single minority group. Instead, in our

[12]This is true even after controlling for whether individuals knew anyone else who had participated in the study before. See Model 6 in Table 8 in the Appendix.

design we opted for cleanly separating between contextual/interactional and composition effects, so that we only conducted mixed sessions with Italians and one immigrant minority at a time.

Second, varying the minority share may also be important for cooperation dynamics, as the size and growth of immigrant and minority populations have been linked to perceived threat and greater levels of out-group discrimination (Adida et al., 2016b). Although we have tried to explore the relative size effect varying the number of Moroccans in ethnically mixed sessions from one through three,[13] we have too little statistical power to conduct any meaningful tests that would address this question (i.e. we only have 9 sessions with one Moroccan and five Italian participants).

Third, we did not elicit beliefs about the contributions of others prior to measuring contributions to the public good. However, research has shown that implementing such belief-elicitation tasks can further help to shed light on the question of the extent to which the willingness to contribute to the public good is driven by preferences or beliefs about the behavior of others (Bigoni, Bortolotti, Casari, & Gambetta, 2019). Lastly, we conducted the experiment in an urban context, the Milan metropolis, and future research needs to determine the extent to which our results generalize beyond this specific setting.

As a final point, this article tried to answer the important question of whether diversity deficits in ethnically mixed groups remain when background information on individuals' levels of integration into society are taken into account. While we cannot claim causality because we did not manipulate individuals level of integration, our observational results suggest that integration, particularly through employment, language acquisition and citizenship, can foster greater contributions to the public good. This finding is in line with research showing causal evidence that naturalization is a catalyst for improving the social integration of immigrants (Hainmueller, Hangartner, & Pietrantuono, 2017). Future research could use a similar research design to examine the causal effect of naturalization or specific labor market integration policies on public goods production. For policymakers who are interested in maintaining and ensuring the provision of public goods in diverse communities, this means that providing avenues

[13]see Table 12 in the Appendix for a breakdown of contributions by size of Moroccan participant share in mixed Moroccan-Italian groups

for the successful integration of immigrants may have benefits for the overall level of public goods provision in society.

6. References

References

Abascal, M., & Baldassarri, D. (2015). Love thy neighbor? ethnoracial diversity and trust reexamined. *American Journal of Sociology*, *121*(3), 722–782.

Abascal, M., Ganter, F., & Baldassarri, D. (2021). Greater diversity or fewer whites? disentangling diversity and minority share at macro and micro levels. *Under Review*.

Adida, C. L., Laitin, D. D., & Valfort, M.-A. (2016a). *Why muslim integration fails in christian-heritage societies*. Harvard University Press.

Adida, C. L., Laitin, D. D., & Valfort, M.-A. (2016b). one muslim is enough! evidence from a field experiment in france. *Annals of Economics and Statistics/Annales d'Économie et de Statistique*(121/122), 121–160.

Alesina, A., Baqir, R., & Easterly, W. (1999). Public goods and ethnic divisions. *The Quarterly Journal of Economics*, *114*(4), 1243–1284.

Alesina, A., & Ferrara, E. L. (2005). Ethnic diversity and economic performance. *Journal of economic literature*, *43*(3), 762–800.

Alesina, A., Glaeser, E., & Glaeser, E. L. (2004). *Fighting poverty in the us and europe: A world of difference*. Oxford University Press.

Alesina, A., & La Ferrara, E. (2002). Who trusts others? *Journal of public economics*, *85*(2), 207–234.

Alesina, A., Miano, A., & Stantcheva, S. (2018). *Immigration and redistribution* (Tech. Rep.). National Bureau of Economic Research.

Allport, G. W., Clark, K., & Pettigrew, T. (1954). *The nature of prejudice*. Addison-wesley Reading, MA.

Ambrosini, M. (2013). Immigration in italy: Between economic acceptance and political rejection. *Journal of international migration and integration*, *14*(1), 175–194.

Auspurg, K., Schneck, A., & Hinz, T. (2019). Closed doors everywhere? a meta-analysis of field experiments on ethnic discrimination in rental housing markets. *Journal of Ethnic and Migration Studies*, *45*(1), 95–114.

Bader, F., & Keuschnigg, M. (2020). Bounded solidarity in cross-national encounters:

Individuals share more with others from poor countries but trust them less. *Sociological Science*, *7*, 415–432.

Baldassarri, D. (2009). Collective action. *Handbook of Analytical Sociology*, 391–417.

Baldassarri, D. (2020). Market integration accounts for local variation in generalized altruism in a nationwide lost-letter experiment. *Proceedings of the National Academy of Sciences*, *117*(6), 2858–2863.

Baldassarri, D., & Abascal, M. (2020). Diversity and prosocial behavior. *Science*, *369*(6508), 1183–1187.

Bechtel, M. M., & Scheve, K. F. (2017). Who cooperates? reciprocity and the causal effect of expected cooperation in representative samples. *Journal of Experimental Political Science*, *4*(3), 206–228.

Berge, L. I. O., Bjorvatn, K., Galle, S., Miguel, E., Posner, D. N., Tungodden, B., & Zhang, K. (2020). Ethnically biased? experimental evidence from kenya. *Journal of the European Economic Association*, *18*(1), 134–164.

Bertocchi, G. (2016). The legacies of slavery in and out of africa. *IZA Journal of Migration*, *5*(1), 24.

Bigoni, M., Bortolotti, S., Casari, M., & Gambetta, D. (2019). At the root of the north–south cooperation gap in italy: Preferences or beliefs? *The Economic Journal*, *129*(619), 1139–1152.

Blum, A., Hazlett, C., & Posner, D. N. (2020). Measuring ethnic bias: Can misattribution-based tools from social psychology reveal group biases that economics games cannot? *Political Anlysis*.

Brewer, M. B. (1999). The psychology of prejudice: Ingroup love and outgroup hate? *Journal of social issues*, *55*(3), 429–444.

Castelli Gattinara, P. (2017). The refugee crisis in italy as a crisis of legitimacy. *Contemporary Italian Politics*, *9*(3), 318–331.

Cettolin, E., & Suetens, S. (2019). Return on trust is lower for immigrants. *The Economic Journal*, *129*(621), 1992–2009.

Chen, D. L., Schonger, M., & Wickens, C. (2016). otreean open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97.

Choi, D. D., Poertner, M., & Sambanis, N. (2019). Parochialism, social norms, and discrimination against immigrants. *Proceedings of the National Academy of Sciences*, *116*(33), 16274–16279.

Choi, J.-K., & Bowles, S. (2007). The coevolution of parochial altruism and war. *science*, *318*(5850), 636–640.

Crowder, K., Hall, M., & Tolnay, S. E. (2011). Neighborhood immigration and native out-migration. *American sociological review*, *76*(1), 25–47.

Del Boca, D., & Venturini, A. (2016). Migration in italy is backing the old age welfare. In *Labor migration, eu enlargement, and the great recession* (pp. 59–83). Springer.

Dinesen, P. T., Schaeffer, M., & Sønderskov, K. M. (2020). Ethnic diversity and social trust: A narrative and meta-analytical review. *Annual Review of Political Science*, *23*.

Dinesen, P. T., & Sønderskov, K. M. (2015). Ethnic diversity and social trust: Evidence from the micro-context. *American Sociological Review*, *80*(3), 550–573.

DAgostino, A., Regoli, A., Cornelio, G., & Berti, F. (2016). Studying income inequality of immigrant communities in italy. *Social Indicators Research*, *127*(1), 83–100.

Enos, R. D. (2017). *The space between us: Social geography and politics*. Cambridge University Press.

Enos, R. D., & Gidron, N. (2016). Intergroup behavioral strategies as contextually determined: Experimental evidence from israel. *The Journal of Politics*, *78*(3), 851–867.

Enos, R. D., & Gidron, N. (2018). Exclusion and cooperation in diverse societies: Experimental evidence from israel. *American Political Science Review*, *112*(4), 742–757.

Fershtman, C., & Gneezy, U. (2001). Discrimination in a segmented society: An experimental approach. *The Quarterly Journal of Economics*, *116*(1), 351–377.

Gereke, J., Schaub, M., & Baldassarri, D. (2018). Ethnic diversity, poverty and social trust in germany: Evidence from a behavioral measure of trust. *PloS one*, *13*(7), e0199834.

Gereke, J., Schaub, M., & Baldassarri, D. (2020). Gendered discrimination against

immigrants: experimental evidence. *Frontiers in Sociology*, *5*, 59.

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

Hainmueller, J., Hangartner, D., & Pietrantuono, G. (2017). Catalyst or crown: Does naturalization promote the long-term social integration of immigrants? *American Political Science Review*, *111*(2), 256–276.

Henrich, J. P., Boyd, R., Bowles, S., Fehr, E., Camerer, C., Gintis, H., et al. (2004). *Foundations of human sociality: Economic experiments and ethnographic evidence from fifteen small-scale societies*. Oxford University Press on Demand.

Istat. (2017). *Non-eu citizens: presence, new inflows and acquisition of citizenship* (Tech. Rep.).

Keuschnigg, M., & Schikora, J. (2014). The dark side of leadership: An experiment on religious heterogeneity and cooperation in india. *The Journal of Socio-Economics*, *48*, 19–26.

Kustov, A., & Pardelli, G. (2018). Ethnoracial homogeneity and public outcomes: The (non) effects of diversity. *American Political Science Review*, *112*(4), 1096–1103.

Laurence, J., & Bentley, L. (2016). Does ethnic diversity have a negative effect on attitudes towards the community? a longitudinal analysis of the causal claims within the ethnic diversity and social cohesion debate. *European Sociological Review*, *32*(1), 54–67.

Legewie, J., & Schaeffer, M. (2016). Contested boundaries: Explaining where ethnoracial diversity provokes neighborhood conflict. *American Journal of Sociology*, *122*(1), 125–161.

Marshall, L., & Paler, L. (2020). Class, ethnicity, and cooperation among women: Evidence from a public goods experiment in lebanon. *Journal of Experimental Political Science*, 1–15.

Miguel, E., & Gugerty, M. K. (2005). Ethnic diversity, social sanctions, and public goods in kenya. *Journal of public Economics*, *89*(11-12), 2325–2368.

OECD. (2019). *Stocks of foreign-born population in oecd countries (indicator)* (Tech. Rep.). Author. (Accessed on 29 October 2019)

Parreñas, R. S. (2016). Perpetually foreign: Filipina migrant domestic workers in rome. In *Migration and domestic work* (pp. 111–124). Routledge.

Pettigrew, T. F., & Tropp, L. R. (2013). *When groups meet: The dynamics of intergroup contact*. psychology press.

Portes, A., & Vickstrom, E. (2011). Diversity, social capital, and cohesion. *Annual review of sociology*, *37*, 461–479.

Putnam, R. D. (2007). E pluribus unum: Diversity and community in the twenty-first century the 2006 johan skytte prize lecture. *Scandinavian political studies*, *30*(2), 137–174.

Sampson, R. J. (2012). *Great american city: Chicago and the enduring neighborhood effect*. University of Chicago Press.

Schaub, M. (2017, October). Threat and parochialism in intergroup relations: Lab-in-the-field evidence from rural Georgia. *Proceedings of the Royal Society B*, *284*(1865), 20171560.

Schaub, M., Gereke, J., & Baldassarri, D. (2020). Does poverty undermine cooperation in multiethnic settings? evidence from a cooperative investment experiment. *Journal of Experimental Political Science*, *7*(1), 27–40.

Sidanius, J., & Pratto, F. (2001). *Social dominance: An intergroup theory of social hierarchy and oppression*. Cambridge University Press.

Sniderman, P. M., Peri, P., de Figueiredo Jr, R. J., & Piazza, T. (2002). *The outsider: Prejudice and politics in italy*. Princeton University Press.

Stichnoth, H., & Van der Straeten, K. (2013). Ethnic diversity, public spending, and individual support for the welfare state: A review of the empirical literature. *Journal of Economic Surveys*, *27*(2), 364–389.

Sturgis, P., Brunton-Smith, I., Read, S., & Allum, N. (2011). Does ethnic diversity erode trust? putnam's' hunkering down'thesis reconsidered. *British journal of political science*, 57–82.

Tajfel, H., Turner, J. C., Austin, W. G., & Worchel, S. (1979). An integrative theory of intergroup conflict. *Organizational identity: A reader*, *56*, 65.

Tusicisny, A. (2017). Reciprocity and discrimination: An experiment of hindu-muslim cooperation in indian slums. *Political Psychology*, *38*(3), 409–426.

Venturini, A., & Villosio, C. (2018). Are migrants an asset in recession? insights from italy. *Journal of Ethnic and Migration Studies*, *44*(14), 2340–2357.

Winter, F., & Zhang, N. (2018). Social norm enforcement in ethnically diverse communities. *Proceedings of the National Academy of Sciences*, *115*(11), 2722–2727.

WITHHELD, C. (2021).

Working Paper.

Zhang, N., Aidenberger, A., Rauhut, H., & Winter, F. (2019). Prosocial behaviour in interethnic encounters: Evidence from a field experiment with high-and low-status immigrants. *European Sociological Review*, *35*(4), 582–597.

Zhang, N., Gereke, J., & Baldassarri, D. (2020). Everyday discrimination in public spaces: A field experiment in the milan metro. *Working Paper*, *XXX*(X), XXX.

7. Appendix

7.1. *Additional Details on the Sampling of Study Participants*

7.2. *Additional Details on the Sampling of Study Participants*

Three criteria were used to select Milanese neighborhoods to sample from: ethnic diversity, wealth, and urbanity. For ethnic diversity, we examined the distribution of the percentage of foreigners in NILs and selected a reasonable interval around the mean (i.e. 25-75 percentile). We excluded 11 neighborhoods with low diversity between 0 -5% and 4 neighborhoods with high diversity levels 28-55%.

Second, we wanted to select participants from neighborhoods with relatively similar levels of wealth. The best proxy is the local price of real estate (cost per m2). Again, we looked at the average size (mean= 85m2, min=71m2, max, 126m2) and cost for private usage of property in diverse neighborhoods (as selected above). In addition, we took into account unemployment as another indicator of wealth. The mean unemployment rate across the 83 NILs was 6.9 % with a min of 4.1% and a max of 11.8%. For each of the wealth indicators we cut the tails at 5%. Lastly, we considered population density as we aimed to include neighborhoods with an urban landscape. We thus excluded those with very small population density (unless this is due to a park, train station etc).

We hired a well-known survey company in Milan to conduct the recruitment of participants from these neighborhoods. For Italian participants, they could also sample from their existing access panel. The recruitment of our immigrant participants was more difficult and also relied in the end on contacting ethnic community centers (e.g. Filipino churches) and a snowball sampling technique. However, we tried to ensure that participants within the same session did not know each other and also included additional questions of whether participants knew anyone who had previously participated in the study in our post-experimental survey at the end of each session (see full summary statistics).

7.3. *Descriptive and Balance Tables*

We estimate all results in the main text with and without control variables. The demographic control variables come from the self-administered survey that was completed before the start of the public goods experiment. The implementation controls that account for game understanding, whether participants knew any other participant who previously participated in the study and the their social networks come from a self-administered survey that was completed after the public goods experiment. We include these control variables in our regression models to improve precision, and control for sample imbalances (see Table 4) as there could be a number of confounding factors associated with ethnicity.

Table 3. Summary Statistics of all Control Variables

	Italian					Filipino					Moroccan				
	Mean	SD	Min	Max	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max	N
Age	40.95	13.56	18	67	561	37.19	13.60	18	64	140	39.39	12.36	18	67	193
Male	0.46	0.50	0	1	561	0.39	0.49	0	1	140	0.74	0.44	0	1	193
Education:															
Primary	0.00	0.04	0	1	561	0.01	0.08	0	1	140	0.20	0.40	0	1	193
Middle	0.04	0.19	0	1	561	0.22	0.42	0	1	140	0.31	0.46	0	1	193
Highschool / Vocational	0.64	0.48	0	1	561	0.47	0.50	0	1	140	0.37	0.48	0	1	193
University	0.32	0.47	0	1	561	0.30	0.46	0	1	140	0.12	0.32	0	1	193
Num. PGG comprehension questions correctly answered	2.46	0.80	0	3	561	1.72	0.95	0	3	140	1.70	0.96	0	3	193
Know another participant?	0.07	0.25	0	1	556	0.29	0.45	0	1	140	0.23	0.42	0	1	192
Employed	0.70	0.46	0	1	561	0.62	0.49	0	1	140	0.23	0.42	0	1	193
Fluent in Italian						0.22	0.41	0	1	125	0.34	0.47	0	1	176
Italian citizen						0.15	0.36	0	1	125	0.19	0.40	0	1	176
5+ Italian Friends						0.51	0.50	0	1	125	0.58	0.50	0	1	176
10+ Italian Acquaintances						0.61	0.49	0	1	125	0.70	0.46	0	1	176

Table 4. Balance Checks

	Italian	Filipino	Moroccan	p-values:	
	(1)	(2)	(3)	(1) vs (2)	(1) vs (3)
A. Homogenous Sessions					
Age	42.1	35.6	41.5	0.000***	0.713
Male	0.48	0.43	0.86	0.429	0.000***
Education:					
Primary	0.00	0.01	0.19	0.097	0.000***
Middle	0.02	0.19	0.38	0.000***	0.000***
Highschool / Vocational	0.64	0.46	0.38	0.007**	0.000***
University	0.34	0.33	0.06	0.938	0.000***
Know another participant?	0.03	0.25	0.19	0.000***	0.000***
Employed	0.74	0.63	0.19	0.060	0.000***
Num. PGG comprehension questions correctly answered	2.44	1.74	1.90	0.000***	0.000***
B. Mixed Sessions					
Age	40.3	38.9	38.2	0.440	0.125
Male	0.45	0.34	0.67	0.083	0.000***
Education:					
Primary	0.00	0.00	0.21	0.665	0.000***
Middle	0.05	0.25	0.27	0.000***	0.000***
Highschool Vocational	0.64	0.49	0.36	0.017*	0.000***
University	0.31	0.26	0.16	0.443	0.001**
Know another participant?	0.09	0.32	0.25	0.000***	0.000***
Employed	0.68	0.62	0.25	0.290	0.000***
Num. PGG comprehension questions correctly answered	2.47	1.71	1.59	0.000***	0.000***
p-values: Homogenous vs Mixed					
Age	0.131	0.143	0.073		
Male	0.453	0.262	0.003**		
Education:					
Primary	0.460	0.329	0.839		
Middle	0.112	0.429	0.138		
Highschool / Vocational	0.957	0.749	0.874		
University	0.511	0.376	0.035*		
Know another participant?	0.003**	0.336	0.375		
Employed	0.142	0.929	0.392		
Num. PGG comprehension questions correctly answered	0.654	0.852	0.027*		

Note: Table reports balance tests using data from the self-administered survey. P-values are from a two-tailed test * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

7.4. *Full Models of Results Reported in the Main Text*

Table 5. Contributions to the Public Good by Homogeneous IT and Mixed Sessions

	(1)	(2)	(3)
Group Composition	0.432	0.426	0.147
	(0.267)	(0.269)	(0.275)
Age		0.002	0.006
		(0.009)	(0.009)
Male		-0.063	0.086
		(0.239)	(0.241)
Education			
Primary School			-1.804**
			(0.681)
Middle School			-0.293
			(0.462)
University			0.392
			(0.268)
Game Understanding			0.287*
			(0.136)
Knowing s.o. who previously participated in the study			-0.325
			(0.385)
Constant	6.301***	6.246***	5.460***
	(0.140)	(0.390)	(0.570)
N	750	750	744

Note: Table reports coefficients, robust standard errors are reported in parentheses (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). Model 1 in Table 5 shows the main result from an OLS regression comparing contributions to the public good in homogeneous Italian groups with those in ethnically mixed groups. Group composition is a dummy variable that takes the values 1 "homogeneous IT group" and 0 "ethnically mixed group". Male is a dummy with values 1 "Male" and 0 "Female". In Model 2, we add demographic controls for age and gender for unequal treatment assignment probabilities across strata. In Model 3, we add further controls for educational attainment, game understanding and personal knowledge of someone who has previously participated in the study. Knowing someone who previously participated in the game is also a dummy that takes the value 1 "Yes" and 0 "No". The reference category for level of completed education is high school or vocational training. The number of observations varies between Models 1 (N=750) and 3 (N=744) because of missing values on some of the control variables from the post-experimental survey used in Model 3.

Table 6. Contributions to the Public Good by Nationality in Homogeneous Sessions

	(1)	(2)	(3)
Filipino	-0.038	-0.070	0.334
	(0.458)	(0.479)	(0.486)
Moroccan	-1.274**	-1.223**	-1.044^{+}
	(0.426)	(0.457)	(0.573)
Age		-0.004	-0.000
		(0.014)	(0.014)
Male		-0.142	-0.332
		(0.376)	(0.389)
Education			
Primary School			-1.106
			(0.696)
Middle School			1.079^{+}
			(0.606)
University			-0.558
			(0.431)
Game Understanding			0.474*
			(0.198)
Knowing s.o. who previously participated in the study			-1.043^{+}
			(0.586)
Constant	6.732***	6.958***	5.946***
	(0.228)	(0.662)	(0.894)
N	342	342	342

Note: Table reports coefficients, robust standard errors are reported in parentheses ($+p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). Model 1 in Table 6 shows the main result from an OLS regression comparing contributions to the public good in homogeneous sessions by nationality of the participants. The baseline comparison group is homogeneous Italian sessions. In Model 2, we add demographic controls for age and gender for unequal treatment assignment probabilities across strata. In Model 3, we add further controls for educational attainment, game understanding and personal knowledge of someone who has previously already participated in the study. Knowing someone who previously participated in the game is also a dummy that takes the value 1 "Yes" and 0 "No". The reference category for level of education is completion of high school or vocational training.

Table 7. Contributions to the Public Good by All Homogeneous and Mixed Sessions

	(1)	(2)	(3)
Group Composition	0.155	0.170	0.113
	(0.224)	(0.225)	(0.222)
Age		-0.002	0.001
		(0.008)	(0.009)
Male		-0.197	-0.007
		(0.219)	(0.224)
Education			
Primary School			-1.842***
			(0.488)
Middle School			0.046
			(0.368)
University			0.344
			(0.254)
Game Understanding			0.335**
			(0.119)
Knowing s.o. who participated in the study			-0.451
			(0.328)
Constant	6.301***	6.471***	5.589***
	(0.140)	(0.364)	(0.510)
N	894	894	888

Note: Table reports coefficients, robust standard errors are reported in parentheses (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). Model 1 in Table 7 shows the main result from an OLS regression comparing contributions to the public good in all homogeneous groups with those in ethnically mixed groups. Group composition is a dummy variable that takes the values 1 "homogeneous groups" and 0 "ethnically mixed groups". Male is a dummy with values 1 "Male" and 0 "Female". In Model 2, we add demographic controls for age and gender for unequal treatment assignment probabilities across strata. In Model 3, we add further controls for educational attainment, game understanding and personal knowledge of someone who has previously participated in the study. Knowing someone who previously participated in the game is also a dummy that takes the value 1 "Yes" and 0 "No". The reference category for level of completed education is high school or vocational training. The number of observations varies between Models 1 (N=894) and 3 (N=888) because of missing values on some of the control variables from the post-experimental survey used in Model 3.

Table 8. Public Goods Contributions by Nationality and Group Composition

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Group Composition	0.096	0.057	0.013	0.959^{+}	0.897	0.958	-0.153	0.193	-0.031
	(0.280)	(0.280)	(0.282)	(0.578)	(0.592)	(0.596)	(0.486)	(0.501)	(0.513)
Age		0.016	0.019^{+}		-0.014	-0.000		-0.046*	-0.046*
		(0.010)	(0.010)		(0.022)	(0.026)		(0.022)	(0.022)
Male		0.322	0.366		0.140	0.101		-1.021^{+}	-0.735
		(0.264)	(0.268)		(0.598)	(0.604)		(0.617)	(0.638)
Education									
Primary School			-2.260***			3.701***			-1.400*
			(0.598)			(0.780)			(0.604)
Middle School			0.069			1.268^{+}			-0.038
			(0.674)			(0.722)			(0.573)
University			0.543^{+}			0.349			0.501
			(0.296)			(0.787)			(0.886)
Game Understanding			0.139			0.672^{+}			0.363
			(0.164)			(0.345)			(0.256)
Knowing s.o. who previously participated in the study			-0.034			0.456			-1.187*
			(0.588)			(0.627)			(0.545)
Constant	6.636***	5.852***	5.207***	5.735***	6.251***	4.001**	5.612***	8.041***	7.831***
	(0.163)	(0.443)	(0.684)	(0.419)	(1.013)	(1.336)	(0.326)	(0.809)	(1.002)
N	561	561	556	140	140	140	193	193	192

Note: Table reports coefficients, robust standard errors are reported in parentheses ($+p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). Table 8 shows the main result from an OLS regression comparing contributions to the public good by nationality and group composition. Models 1-3 show the results for Italians. Models 4-6 show the results for Filipino participants, and Models 7-9 for Moroccan participants. Group composition is a dummy variable that takes the values 1 "homogeneous group" and 0 "ethnically mixed group". In Models 2, 5 and 8, we add demographic controls for age and gender for unequal treatment assignment probabilities across strata. In Models 3, 6 and 9 we add further controls for educational attainment, game understanding and personal knowledge of someone who has previously already participated in the study. Knowing someone who previously participated in the game is coded as a dummy that takes the value 1 "Yes" and 0 "No". The reference category for level of education is completion of high school or vocational training.

Table 9. Contributions to the Public Good by Italians in Mixed Sessions by Nationality of Immigrants

	(1)	(2)	(3)	(4)	(5)
Moroccan Mixed Group Composition	-0.679^{+}	-0.660^{+}	-0.678^{+}	-0.743^{+}	-0.124
	(0.367)	(0.363)	(0.358)	(0.377)	(0.585)
Age		0.021	0.030*	0.029*	0.028^{+}
		(0.014)	(0.015)	(0.015)	(0.015)
Male		0.437	0.500	0.465	0.464
		(0.363)	(0.358)	(0.373)	(0.375)
Primary School			-2.261**	-2.498**	-2.330**
			(0.779)	(0.811)	(0.812)
Middle School			-0.492	-0.472	-0.514
			(0.952)	(0.898)	(0.921)
University			1.296***	1.329***	1.238**
			(0.384)	(0.388)	(0.394)
Game Understanding			0.109	0.121	0.102
			(0.226)	(0.226)	(0.230)
Knowing s.o. who previously participated in the study			-0.152	-0.205	-0.157
			(0.731)	(0.731)	(0.733)
1 male immigrant				-0.656	-0.389
				(0.466)	(0.591)
2 male immigrants				-0.205	0.778
				(0.459)	(0.652)
1 Moroccan male					-0.669
					(0.976)
2 Moroccan male					-1.534^{+}
					(0.866)
Constant	7.029***	5.952***	4.942***	5.226***	5.107***
	(0.253)	(0.612)	(0.990)	(1.031)	(1.036)
N	276	276	276	276	276

Note: Table reports coefficients, robust standard errors are reported in parentheses ($+p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). Model 1 in Table 9 shows the main result from an OLS regression comparing contributions to the public good by Italians in mixed groups with two immigrants and four Italians by nationality of the immigrant participants. Group composition is a dummy variable with the reference category "mixed Filipino-Italian groups". Male is a dummy with values 1 "Male" and 0 "Female". In Model 2, we add demographic controls for age and gender for unequal treatment assignment probabilities across strata. In Model 3, we add further controls for educational attainment, game understanding and personal knowledge of someone who has previously participated in the study. Knowing someone who previously participated in the game is also a dummy that takes the value 1 "Yes" and 0 "No". The reference category for level of completed education is high school or vocational training. In Model 4, we add a control for the number of immigrant men in each session, ranging from 0 to max. 2. In Model 5, we further add an interaction term with nationality of immigrants and the number of male immigrant participants in the group.

Table 10. Public Goods Contributions by Employment Status for Italians and Immigrants

	(1)	(2)	(3)
IT not employed	0.880*	0.858*	0.482
	(0.345)	(0.349)	(0.388)
IT employed	1.220***	1.260***	0.773*
	(0.287)	(0.297)	(0.367)
IMM employed	0.689^{+}	0.708^{+}	0.567
	(0.382)	(0.389)	(0.402)
Homogeneous Group Composition	0.202	0.212	0.139
	(0.224)	(0.224)	(0.225)
Age		-0.008	-0.004
		(0.008)	(0.009)
Male		-0.058	0.024
		(0.222)	(0.226)
Education			
Primary School			-1.424**
			(0.523)
Middle School			0.308
			(0.384)
University			0.271
			(0.258)
Game Understanding			0.273*
			(0.128)
Knowing s.o. who previously participated in the study			-0.348
			(0.344)
Constant	5.479***	5.816***	5.336***
	(0.263)	(0.414)	(0.538)
N	894	894	888

Note: Table reports coefficients, robust standard errors are reported in parentheses ($+p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). Table 10 shows the result from an OLS regression comparing contributions to the public good by employment level of immigrants and Italians controlling for their assignment to either ethnically homogeneous or mixed groups (dummy variable, reference category is assignment to mixed groups). Model 1 shows the main results with the reference category being immigrants who are currently not active in the labor market (i.e. unemployed, retired, housewives, maternity leave). In Model 2 we add demographic controls for age and gender for unequal treatment assignment probabilities across strata. In Model 3, we add further controls for educational attainment, game understanding and personal knowledge of someone who has previously already participated in the study. Knowing someone who previously participated in the game is coded as a dummy that takes the value 1 "Yes" and 0 "No". The reference category for level of education is completion of high school or vocational training. The Number of observations varies between Models 1 (N=894) and 3 (N=888) because of missing values on some of the control variables from the post-experimental survey used in Model 3.

Table 11. Public Goods Contributions of Immigrants by Level of Integration

	(1)	(2)	(3)	(4)
No IT Citizenship	1.347*			
	(0.529)			
No Fluency in Italian		-0.875^{+}		
		(0.463)		
Fewer than 5 IT friends			-1.006*	
			(0.395)	
Fewer than 10 IT acquaintances				-0.895*
				(0.404)
Homogeneous Group Composition	0.838*	0.736^{+}	0.676^{+}	0.635
	(0.400)	(0.391)	(0.391)	(0.392)
_cons	5.124***	6.031***	5.884***	5.751***
	(0.301)	(0.439)	(0.322)	(0.315)
N	301	301	301	301

Note: Table reports coefficients, robust standard errors are reported in parentheses ($+p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). Table 11 shows the result from an OLS regression comparing contributions to the public good by immigrants integration level controlling for their assignment to either ethnically homogeneous or mixed groups (dummy variable, reference category is assignment to mixed groups). Model 1 shows contributions levels for immigrants with and without Italian citizenship. Model 2 shows results comparing contributions of immigrants who self-report being fluent in Italian compared to those who indicate non-fluency in Italian. In Model 3, we compare contributions of immigrants who are socially embedded in terms of friendships (5 or more Italian friends) with those who have fewer than 5 Italian friends. Model 4, similarly compares immigrants who report having 10 or more Italian acquaintances with those who report to have fewer than 10 Italian acquaintances.

Table 12. Public Goods Contributions by Immigrant Share in Mixed Moroccan-Italian Groups

	(1)	(2)	(3)
Mc1mixed	1.025*	0.931^{+}	1.074*
	(0.483)	(0.492)	(0.528)
MC3mixed	0.182	0.166	0.305
	(0.442)	(0.441)	(0.439)
Moroccan	-0.692^{+}	-0.858*	0.093
	(0.396)	(0.412)	(0.520)
Age		-0.020	-0.012
		(0.014)	(0.015)
Male		0.455	0.551
		(0.366)	(0.370)
Education			
Primary School			-1.513^{+}
			(0.813)
Middle School			-0.270
			(0.636)
University			0.751^{+}
			(0.432)
Game Understanding			0.461*
			(0.218)
Knowing s.o. who previously participated in the study			-0.373
			(0.556)
Constant	6.164***	6.816***	5.063***
	(0.252)	(0.632)	(0.958)
N	348	348	342

Note: Table reports coefficients, robust standard errors are reported in parentheses ($+p < 0.10$,, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). Table 12 shows the result from an OLS regression comparing contributions to the public good by the share of Moroccan participants in mixed Italian-Moroccan groups. The baseline category are groups with two Moroccans and four Italians (same immigrant share as in all mixed Filipino-Italian groups). Model 1 shows the main result for Moroccan share ranging from one to three Moroccan participants out of a total of six controlling only for contributions made by Moroccan participants. In Model 2, we add demographic controls for age and gender for unequal treatment assignment probabilities across strata. In Model 3, we add further controls for educational attainment, game understanding and personal knowledge of someone who has previously already participated in the study. Knowing someone who previously participated in the game is coded as a dummy that takes the value 1 "Yes" and 0 "No". The reference category for level of education is completion of high school or vocational training.