

**Cooperation in a Fragmented
Society: Experimental
Evidence on Syrian Refugees
and Natives in Lebanon**

Michalis Drouvelis, Bilal Malaeb, Michael Vlassopoulos, Jackline Wahba

Impressum:

CESifo Working Papers

ISSN 2364-1428 (electronic version)

Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH

The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute

Poschingerstr. 5, 81679 Munich, Germany

Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de

Editor: Clemens Fuest

www.cesifo-group.org/wp

An electronic version of the paper may be downloaded

- from the SSRN website: www.SSRN.com
- from the RePEc website: www.RePEc.org
- from the CESifo website: www.CESifo-group.org/wp

Cooperation in a Fragmented Society: Experimental Evidence on Syrian Refugees and Natives in Lebanon

Abstract

Lebanon is the country with the highest density of refugees in the world, raising the question of whether the host and refugee populations can cooperate harmoniously. We conduct a lab-in-the-field experiment in Lebanon studying intra- and inter-group behavior of Syrian refugees and Lebanese nationals in a repeated public good game without and with punishment. We find that homogeneous groups, on average, contribute and punish significantly more than mixed groups. These patterns are driven by the Lebanese participants. Our findings suggest that it is equally important to provide adequate help to the host communities to alleviate any economic and social pressures.

JEL-Codes: D910, J500, F220.

Keywords: refugees, public good game, cooperation, punishment.

Michalis Drouvelis
Department of Economics
University of Birmingham / UK
m.drouvelis@bham.ac.uk

Bilal Malaeb
Institute of Global Affairs
London School of Economics and Political
Science / London / UK
b.malaeb@lse.ac.uk

Michael Vlassopoulos
Department of Economics
University of Southampton / UK
m.vlassopoulos@soton.ac.uk

Jackline Wahba
Department of Economics
University of Southampton / UK
j.wahba@soton.ac.uk

December 2019

Financial support from the Universities of Birmingham and Southampton is gratefully acknowledged. We are grateful for the field assistance of Nayla Abdel Salam, Raya Abou Jaber, Nagham Al Banna, and Yasser Bohsas (who sadly passed away in late 2018), and for the research assistance of Sofia Gallo. The study was approved by the University of Southampton IRB under Ethics ID:30483.

1. Introduction

In 2018, the global population of forcibly displaced people stood at a record high 70.8 million, according to the UN refugee agency (UNHCR, 2018), representing one of the most pressing global humanitarian challenges of our time. At the end of the same year, Syrians continued to be the largest forcibly displaced population, with 13 million people displaced, including 6.7 million refugees, the majority of whom are hosted by neighboring countries. On the receiving end, Lebanon is one of the largest embracers of refugees, hosting almost one million Syrians in 2018, making it the country with the highest density of refugees in the world. Although Syrians and Lebanese are not ethnically different, and share the same language and culture, the relationship between these two nations has been very strained in the recent past.¹ As a result of the past problems and the recent sentiment that Syrians are to blame for Lebanon's economic woes, Syrians now living as refugees in Lebanon might be experiencing a less welcoming environment.² Hence, whether these two populations can co-exist and co-operate is an intriguing case to study, given that many other populations become displaced and hosted in unwelcoming communities.

Cooperation is fundamental for the provision of public goods. As already well documented in the literature, ethnic diversity can lead to lack of cooperation and therefore less provision of public goods (Alesina and La Ferrara, 2005; Miguel and Gugerty, 2005). Studies show that among social groups, there is a strong tendency to favor one's in-group (e.g. Bernhard, Fischbacher and Fehr, 2006). Interestingly though, little is known about cooperation and reciprocity between refugees and the host community. Refugees differ in several aspects from typical economic migrants, in terms of their demographic characteristics, skills and their motivations for leaving from their home country and for establishing a permanent residence in the receiving country.³ They also represent a particularly vulnerable group due to the psychological trauma associated with exposure to violence and conflict, which might impact altruistic behavior (see Voors et al. 2012, Bauer et al., 2016). On the other side, from the

¹ The historical conflict includes Syrian occupation of Lebanon for 29 years between 1976-2005, accusations of Syrian intervention in Lebanese politics, and suspicions of Syria's involvement in the assassination of Lebanese politicians such as former prime minister, Rafiq Hariri, in 2005. However, in 2008, Syria officially recognized Lebanon's sovereignty and diplomatic relationship was established between the two countries.

² Economist, "Politicians are stoking anti-refugee sentiment in Lebanon" (Aug. 22nd, 2019).

³ Becker and Ferrara (2019) provide a recent survey of the growing literature that focuses on the consequences of forced migration on receiving populations, on migrants themselves and on sending populations for economic, political and other outcomes.

perspective of countries that receive refugees, the reasons for hosting them are primarily humanitarian and not based on possible economic benefits, as in the case of hosting immigrant workers. Yet, there are concerns about the (perceived) burden of hosting refugees, and in our particular case, previous tensions between the two nations might also affect reciprocity and cooperation of the host community.

Against this background, the objective of this paper is to shed empirical light on the intra- and inter-group cooperation of Syrian refugees and the native Lebanese communities. To this end, we study a social dilemma situation in which a conflict emerges between personal and collective interest. This incentive structure characterizes a number of real-life settings such as teamwork, participation in collective actions and provision of public goods, tax compliance, environmental protection and donations to charities. It is also a particularly suitable paradigm to apply in a context where tensions have been built up between two natural groups (further described in section 2), the institutional framework is weak, and refugees are not secluded in refugee camps but are embedded in the wider community.

We carried out a lab-in-the-field experiment in Lebanon with Syrian refugees and Lebanese nationals.⁴ To measure cooperation, we make use of a workhorse game in the experimental economics literature on cooperation and punishment: the linear public goods game without and with punishment (Fehr and Gächter, 2000; 2002). Participants were randomly assigned to Lebanese-only, Syrian-only or mixed groups and played six rounds of the public good game in randomly formed pairs. We use subtle differences in spoken Arabic to make salient the composition of the groups' nationality to participants at the beginning of each session, without explicitly invoking nationality in the experimental instructions. Our main interest lies in the levels of cooperation that are sustained across and within the two groups and the extent to which punishment opportunities to enforce cooperation are effective in raising cooperation. We expect that due to enhanced group identification (Tajfel and Turner, 1979), contributions to the public good will be greater in single-nationality groups than in mixed-nationality groups. In a similar vein, we anticipate that punishment behavior will also be treatment dependent but the direction of the effect is an open empirical question. It might be that higher contributions in homogeneous interactions lead to lower punishment; however, if mixed groups care less

⁴ Gneezy and Imas (2016) define a lab-in-the-field study “as one conducted in a naturalistic environment targeting the theoretically relevant population but using a standardized, validated lab paradigm” (p. 3).

about others in their group, then subjects may be less willing to deviate from the selfish strategy leading to less assignment of punishment than in homogeneous groups.

Our findings show that, on average, contributions to the public good are significantly higher in homogeneous groups compared to mixed groups, suggesting a stronger in-group versus out-group cooperation effect. This result is driven by the host (Lebanese participants) who exhibit a stronger tendency to reduce contributions in the mixed treatment. We also find that average earnings are significantly lower in mixed groups compared to homogeneous groups. In the case of the public good game with punishment, we find a substantial degree of antisocial punishment, especially in Lebanese-only groups. Moreover, mixed groups punish significantly less than homogeneous groups, a result that is again driven by Lebanese participants. This suggests that for the Lebanese hosts, although as expected there is greater cooperation and reciprocity toward own group, there is also a lower willingness to punish out-group defectors and even evidence of an inclination to punish in-group cooperators. Taken together, our findings indicate that mixing of the two groups leads to deterioration of the public good and makes individuals worse off and that punishing opportunities are not able to remedy the situation.

Our study is related to several strands of literature. First, our paper contributes to a literature that examines experimentally the role of diversity for the provision of public goods. Probably the most closely related previous studies are Habyarimana et al. (2007) and Alexander and Christia (2011), which find that cooperation in ethnically mixed groups is lower than that in homogeneous groups in studies that were carried out in Ugandan slums, with different ethnic groups, and in Bosnia with Catholic and Muslim participants, respectively. Ruffle and Sosis (2006) find that kibbutz members are more cooperative toward kibbutz members than they are toward city residents, while Castro (2008) carries out a public good game with participants of British and Italian nationality and finds lower contributions in mixed groups than in homogeneous groups.⁵

⁵ Also, related is a strand of experimental literature that studies in-group and out-group trust of naturally occurring groups, such as, Ashkenazi and Sephardi Jews in Israel (Fershtman and Gneezy, 2001), non-Western immigrants and native Dutch in the Netherlands (Cettolin and Suetens, forthcoming), first-generation immigrants and native-born Americans in the United States (Cox and Orman, 2015), immigrant and native youth in Germany (Felfe et al., 2018) and different ethnic groups in Afghanistan (Bartos and Lively, 2018).

In addition, and more importantly, our study relates to the literature that examines the economic impact of forced displacement on the refugees and on host communities. A growing number of papers have focused on the impact of refugees on the labor market, often with mixed results, which are summarized in two recent surveys (Becker and Ferrara, 2019; Verme and Schuettler, 2019). Few studies have focused on the large and sudden forced migration in high-income countries such as in the US, Israel, France, e.g. Card (1990), Borjas (2017), Clemens and Hunt (2019), while others examined the impact of refugees in low-income countries such as in Tanzania, Kenya, and Sudan e.g. Maystadt and Verwimp (2014), Ruiz and Vargas-Silva (2015) and Alix-Garcia et al. (2018). More specific to our setting, a few recent studies have examined the impact of Syrian refugees on the labor market outcomes of natives in Turkey, for example Tumen (2016) and Del Carpio and Wagner (2015), and in Jordan, for example, Fallah, Krafft and Wahba (2019). Another set of studies has studied the impact of the inflow of Syrian refugees on the well-being of the host population and in particular on the impact on education and housing (e.g., Balkan et al., 2018; Tumen, 2019; and Assaad, Ginn and Saleh, 2019) and on consumer prices (Balkan and Tumen, 2016).

We contribute to this literature by providing direct experimental evidence on the degree of cooperation and reciprocity between a host community and a recently displaced population. In particular, unlike the previous experimental studies cited above, which examine cooperation between different ethnic groups, or natives and immigrant groups, in our context there is no ethnic or cultural divide between the two groups, though there exist tensions between the host population and the refugees. All of the above make the context of our study novel. At the same time, our paper is important for policymakers and humanitarian organisations who are interested in the welfare of refugees and social cohesion in the host community. Lack of cooperation between natives and refugees would be a barrier for the efforts of all those who are striving to ensure decent living for the displaced populations in the host community, and an important challenge faced by societies experiencing a large influx of refugees. The successful integration of the newly arrived individuals into the local communities is essential for the well-being of refugees and the sustainability of hosting refugees.

The structure of the rest of the paper is as follows. Section 2 will provide a background to the relationship between Lebanon and Syria prior to the Syrian conflict. In Sections 3 and 4, we describe our experimental design and hypotheses. Section 5 discusses the data, while we present our main experimental results in Section 6. Finally, Section 7 concludes.

2. Background and Context

In this section, we seek to put the current research question into context, by describing briefly the current Syrian war and refugee crisis, and the history of Lebanese-Syrian relations.

The Syrian war erupted in 2011 after a series of protests, which turned violent between civilians, the state, and militant groups. The war caused a significant number of casualties and millions of forcibly displaced individuals within and outside Syria. In Lebanon, Syria's next-door neighbor with a population of just over four million people, estimates show that as of year 2018 there are around one million registered Syrian refugees (UNHCR, 2019), although unofficial estimates are reportedly higher. This made Lebanon the largest refugee-hosting nation on a per-capita basis. As a result, the refugee influx created a significant load on an already-fragile economy and infrastructure, which has led to increased social tensions between refugees and their Lebanese hosts.

A recent report of a nationally representative survey on social tensions in Lebanon shows that nearly all Lebanese respondents agree that Syrian refugees are placing a significant burden on their country's resources (UNDP, 2018). More importantly, the report cites decreasing intercommunal contact between the two groups and that the propensity for negative collective action (e.g. violence and mass evictions) has increased in recent years. Despite the large inflow of foreign aid into the country to support the Syrian refugees, negative attitudes persist both on the popular level and in the political narrative. These tensions can be traced to a general xenophobic trend that is happening globally (e.g. in Europe and United States) but there are elements that make the Lebanese context more peculiar. Firstly, the Syrian and Lebanese people share common culture and language. Secondly, circular migration between the two countries existed for centuries. Indeed, prior to World War I the two states were part of a single larger nation that included Palestine and Jordan as well – termed Greater Syria (or Ottoman Syria). In fact, it had always been said that the Lebanese and Syrians are 'one people in two states', which makes the nature of the relationship between the two groups particularly interesting. Nevertheless, recent history between the two groups (or states) is far from pleasant. Following the eruption of the Lebanese civil war in 1975, an Arab Deterrent Force – consisting almost exclusively of Syrian army – was mandated to restore peace in the country. Despite the war ending in 1990, the Syrian forces remained in Lebanon until 2005, had a strong influence on the governance of the country, and have been accused of vast corruption. In 2005, following the assassination of the Lebanese prime minister, Syrian authorities were accused of plotting the incident and a United Nations resolution was sought to expel the Syrian forces. For years

after this assassination, Lebanon witnessed a number of political assassinations from its anti-Syrian political camp resulting in notably tense relations between the two countries. Therefore, these atrocities had not yet been wiped from the Lebanese collective memory by the time the Syrian war had erupted and its resultant refugee influx occurred. Finally, in addition to this long but tense history between the two groups, Lebanon is argued to be a fragile state with poor infrastructure, limited governance capacity, and a stagnant economy (Malaeb, 2018); all of which inevitably cause tensions within the society, both among the Lebanese people themselves and with ‘others’.

3. Experimental design and procedures

3.1 Framework

Our simple measure of cooperation is centered on a linear social dilemma game without and with punishment opportunities (Fehr and Gächter, 2000; 2002). In both games, subjects are randomly assigned to a two-person group (henceforth, matching-group). The social dilemma game without punishment options consists of one stage, in which each subject is endowed with 10 tokens and has to decide how many of them to keep and how many to contribute to the public good (described as a ‘project’ to subjects). Each token kept increases the own monetary payoff by one experimental currency unit (ECU). Each token contributed to the public good increases the payoff of every group member by 0.75 ECUs. The payoff function from the first stage is given by equation (1).

$$\pi_i^1 = 10 - g_i + 0.75 \cdot (g_i + \sum_{j=1}^2 g_{j \neq i}), \quad (1)$$

where g_i ($0 \leq g_i \leq 10$) denotes the number of tokens contributed to the public good by group member i .

The social dilemma game with punishment options consists of two stages, of which the first one is identical to the above description. In the second stage, subjects can see the profile of contributions of the other member of the matching-group and are given the opportunity to assign costly punishment points to each of the other two group members. Subjects could assign up to 5 punishment points. Each punishment point costs the punisher 1 ECU and the recipient of the punishment 3 ECUs. Thus, the cost-to-impact ratio is 1:3. The total payoff from both stages is computed as follows:

$$\pi_i = \pi_i^1 - \sum_{j \neq i} p_{ij} - 3 \cdot \sum_{j \neq i} p_{ji}, \quad (2)$$

where π_i denotes group member i 's payoff from the first (contribution) stage and the punishment points group member i assigns to group member j .

Conditional on each subject i being motivated to maximize equation (2), the unique subgame perfect equilibrium requires that subjects free-ride completely in the first stage and refrain completely from punishing in the second stage.

3.2 Experimental treatments

Our experimental design consists of three between subjects treatments, which vary whether the participants in a session were only Syrians, only Lebanese or consist of both Syrians and Lebanese. We refer to the resulting treatments as “Only Syrians”, “Only Lebanese” and “Mixed” treatments, respectively.

Subjects played the first three rounds of the public good game without punishment followed by three rounds of the game with punishment. We employed a partners' matching protocol, whereby the group composition remains the same throughout the experiment. This allows us to observe the dynamics of cooperative behavior and how having experience of the no-punishment game affects behavior in the punishment game (in terms of contributions and assignment of sanctions). Furthermore, we provide incentives for subjects in order to elicit their beliefs about other's contribution and own expectations of being punished.

3.3 Recruitment and procedures

We conducted the experiment in the Aley region, governorate of Mount-Lebanon, as it possesses some characteristics that makes it a suitable location for the research at hand. Firstly, Aley is an urban area and among the top 10 populous towns of Lebanon. Secondly, the area had a long history of circular Syrian migration - Syrian workers historically travelled to this area to work in construction and agriculture, as they have strong ties and a large community that helps them find work. Thirdly, the area is sufficiently close (15km) to the capital, Beirut, to be accessible, but far enough away to offer affordable housing. It is also less than 50 km from the Lebanese-Syrian border making it an easy destination for refugees to reach. For these reasons, the area is home to more than 6000 registered Syrian refugees (according UNHCR – this is an underestimate of the refugee population as many are not registered). Finally, the area is known for its mild, or even welcoming, stance at the political level towards refugees.

We advertised our experiment around 10 days before we started the fieldwork in early November 2017. We relied mainly on word-of-mouth and social media as ways to advertise it.⁶ Subjects who were interested signed up by phone or using an online survey that asked them some personal details (e.g. sex, nationality, age, ability to read and write, their time availability, and their contact details). We randomly assigned sessions into two types: homogeneous sessions (Lebanese only or Syrians only) and mixed sessions (approximately equally split between Lebanese and Syrians). For every two mixed sessions, we allocated two homogeneous sessions (one of each nationality). By the end of the experiment, we had around 420 expressions of interest. Around three days before the start of the experiment, we began a random draw of around 25 participants for each session based on their nationality and declared availability (e.g. early afternoon or late afternoon). Participants who could not make it to the session were asked whether they would like to be placed in the subject pool again for future draws.

The experimental team consisted of a research coordinator (a co-author), and four field assistants (two Lebanese and two Syrians). The fieldwork took place in a large hall of a school in Aley. Each day, from Monday to Saturday, two two-hour sessions took place – one at 3pm and another at 6pm. We ran a total of 14 sessions (4 Syrian only, 3 Lebanese only, and 7 Mixed sessions). When participants arrived, they were all asked to sign a consent form and read an information sheet before entering the hall where the experiment was held.⁷ Before participants arrived, the team numbered the tables in the room where the actual experiment took place such that participants could not sit close to each other. Afterwards, they were asked to randomly draw a numbered ticket, which corresponded to their particular seat. The experimental room was large enough to host up to forty people at once (please see Appendix 2 for an image of the room).

Since our treatment relies on participants' awareness of being in a mixed session or homogenous session, we devised an exercise that would make the composition of the group

⁶ The text of the advertisement read as follows:

“A group of researchers from British Universities are undertaking a study around risk and social preferences in Lebanon that aims to understand how these preferences affect economic decisions of individuals and the economy as a whole. The study consists of interactions between different groups of the society that would make certain decisions as well as filling in a survey. So we invite you to participate if you are aged 18 and above. The study's sessions will be held in Aley, every day, from 13th to 24th of November. You only need to attend one session which should last up to two hours. As an appreciation for your participation, you will be remunerated with a minimum of \$5 for showing up, and your total payoffs can be up to \$60. If you would like to participate, please phone the following number or register your interest using the following link.”

⁷ If there was an odd number of participants (e.g. due to no show), we randomly chose one participant offered them the show-up fee and asked them to leave.

salient in a subtle way. The premise of the exercise is - due to the similarity in physical appearance between Lebanese and Syrians - to rely on the most salient difference between the two groups: their accent in Arabic. The exercise consisted of pictures of products (printed on an A4 laminated paper) that individuals would usually see or purchase on a daily basis in their local shops.⁸ Each participant was asked to say aloud the name of the product that they draw from the list and quote its price as they see it in the market. They were explicitly told that there is no correct answer to the question and that the answers to their question would not affect their payoff. The exercise was framed as an icebreaker and a way to observe the awareness of each individual of the local economic conditions.⁹

After this ‘icebreaker’, the participants were handed the instructions of the public good game without punishment, and the coordinator read the instructions aloud.¹⁰ They were then asked to practice how their payoffs are calculated and were explicitly told that their answers would not affect their payoff. The research assistants roamed the hall and ensured every person had understood how their payoffs were calculated before proceeding. After that, we proceeded to the first stage of the public good game without punishment. This stage was repeated for three rounds in total. We then proceeded to the public good game with punishment, which also consisted of three rounds and we followed a similar procedure.

Participants were told that one round out of the six would be randomly chosen by public draw to determine their actual monetary gains. Finally, participants were asked to complete a short questionnaire while we computed their payoffs, and prepared an envelope with their money in cash.

4. Hypotheses

This section formulates behavioral hypotheses that we seek to test in the public good games we conducted. In the following, our hypotheses refer to both the public good game without and with punishment opportunities.

⁸ Examples of products included: apples, banana, onions, oranges, tomato, parsley, potato, bread, rice, hummus, whole chicken, local yoghurt, cheese, potato crisps, local chewing gum, soda, mineral water, etc.

⁹ In conversations with some participants after the pilot session, we asked whether they had felt any particularly negative or positive emotions as part of this exercise, e.g. boredom, feeling silly, confused, embarrassed, finding it funny, etc. They all agreed, individually, that they did not understand the purpose of the exercise, but were generally indifferent to it.

¹⁰ A copy of the experimental instructions can be found in an appendix.

If we assume that individuals are exclusively self-interested, they will always defect (i.e. contribute nothing) in the public goods game, since defection is a dominant strategy. Similarly, a selfish individual would never punish another player as punishment is monetarily costly. However, there is by now a well-established literature (as surveyed in Fehr and Schmidt, 2006) showing that individuals deviate from what standard economic theory – assuming selfishness and rationality – would predict. Related to our public good game environments, an expansive literature in economics has shown that individuals, to some extent, are willing to cooperate and engage in costly punishment activities (for an overview, see Chaudhuri, 2011).

To understand the role of group membership for intergroup relations and behavior, we draw on the social identity theory (Tajfel and Turner, 1979), which has spawned a large literature in psychology and sociology. In economics, social identity was introduced and formalized more recently by Akerlof and Kranton (2000). Experimental evidence indicates that individuals have group-specific preferences in their cooperative behavior when either “minimal groups” (e.g., Eckel and Grossman, 2005; Chakravarty and Fonseca, 2014) or naturally occurring social groups (e.g., Goette et al., 2012) are considered. Specifically, it has been shown that groups with common identities exhibit higher cooperation levels compared to groups consisting of members with fragmented, heterogeneous identities (Ruffle and Sosis, 2006; Habyarimana et al., 2007; Goette et al., 2006; Castro, 2008; Alexander and Christia, 2011). This leads us to formulate our first hypothesis.

Hypothesis 1. Homogeneous groups contribute more than mixed groups.

The assignment of punishment will, in turn, be affected across homogeneous and mixed groups as a result of their cooperative attitudes. The behavioral effects from meting out punishment are, however, less clear. On the one hand, being more cooperative in homogeneous groups might make the use of punishment less necessary since the enforcement of a cooperation norm has already been established (*Hypothesis 2a*). This hypothesis mirrors in-group favoritism (that is, homogeneous punish less harsh own group members, following their higher cooperation levels). Similarly, lower contribution levels for outgroups may lead to harsher punishment, assuming that outgroups care about enforcing a social norm for high contributions. Alternatively, if this is not the case, members of fragmented social groups may engage less in punishment activities in a way such that they express lack of interest to establish cooperation norms (following lower cooperation levels among mixed groups). This tendency will result in more egoistic behavior and higher willingness to defect, leading to less harsh punishment

assignment in between-group interactions (*Hypothesis 2b*). These arguments are summarized in the following set of two hypotheses.

Hypothesis 2a. Homogeneous groups punish less harsh than mixed groups.

Hypothesis 2b. Homogeneous groups punish harsher than mixed groups.

As a final step, we derive hypotheses relating to groups' welfare as measured by their average net earnings from the public good games. Our expectation is that homogeneous groups which exhibit higher cooperation rates (as specified in *Hypothesis 1*) will have higher earnings from the contribution phase of the public good game. However, when we consider the punishment phase of the game, the extent to which higher welfare is observed between homogeneous and mixed groups depends on their respective attitudes towards using punishment. Existing experimental evidence shows that the assignment of costly punishment is detrimental in individuals' welfare, especially in the short run (e.g., Gächter et al., 2008). Since, in our environment, the assignment of punishment is monetarily costly both for the person who assigns and receives punishment points, we anticipate that if homogeneous groups assign significantly less harsh punishment than mixed groups (following *Hypothesis 2a*), then outgroups' welfare will be lower than in-groups' welfare. In contrast, if homogeneous groups assign significantly harsher punishment than mixed groups (following *Hypothesis 2b*), outgroups will earn more than in-groups, assuming the cost of higher punishment dominates the gain of higher contribution for the homogeneous group. We summarize our last set of two hypotheses as follows.

Hypothesis 3a. Homogeneous groups earn more than mixed groups.

Hypothesis 3b. Homogeneous groups earn less than mixed groups.

5. Data Description

Our sample consists of 312 participants, 78 of whom were randomly assigned to Syrian-only sessions, 70 to Lebanese-only sessions, and 164 to Mixed sessions. Out of the 312 participants, 32 did not complete the entire end-of-experiment survey that collects data on their demographic characteristics, as well as risk and trust preferences. As such, the regression analysis below is

based on the 280 participants that completed the survey, while for the non-parametric tests we use the full sample.¹¹

Table 1 shows the descriptive statistics of the survey variables based on the sample of 280 participants (67 participants - Syrian-only sessions; 66 participants - Lebanese-only sessions; 147 participants - Mixed sessions). As can be seen, participants are on average about 30 years of age, while the gender composition is balanced. With regard to education, there are some notable differences across the two groups, with the Lebanese participants being more likely to have acquired University education than the Syrians. Also, in terms of religion, Lebanese participants are mostly Druze, whereas the Syrians are mostly Sunni.

We performed balance tests to check whether Lebanese participants in mixed sessions were similar in observable characteristics to Lebanese participants in homogeneous sessions, and similarly for Syrian Participants. To do that, we run regressions for each of the descriptive variables (in Table 1) on the treatment assignment, for each nationality separately. We find that the groups were balanced for all observable characteristics with some exceptions. We find that Syrian participants had higher proportions of having intermediate education and non-university diploma in mixed groups compared to Syrians in homogenous sessions. Lebanese participants had slightly more males and were more prone to take risk in mixed groups than in homogenous ones. Not surprisingly, we also find that both Lebanese and Syrian participants knew approximately half as many participants in the mixed sessions than in homogeneous sessions. This can be attributed to having half as many people from the same nationality in a mixed group session. Together, these results suggest that the randomization to the treatment was well balanced.

¹¹ Results of the regressions with the full sample of 312 participants and without including control variables derived from the survey are presented in Appendix 1.

Table 1. Descriptive statistics across treatments

| Variable | Homogenous | | Mixed | | |
|------------------------------|---------------------|----------------------|----------------------|----------------------|----------------------|
| | Syrian Only (1) | Lebanese Only (2) | All (3) | Syrian (4) | Lebanese (5) |
| Age | 29.58 (9.62) | 27.18 (11.15) | 30.13 (11.26) | 31.23 (11.17) | 28.82 (11.31) |
| Male | 0.45 (0.50) | 0.45 (0.50) | 0.54 (0.50) | 0.59* (0.50) | 0.49 (0.50) |
| Married | 0.33 (0.47) | 0.76 (0.43) | 0.53 (0.50) | 0.31 (0.47) | 0.79 (0.41) |
| No. of Children | 1.60 (1.81) | 0.39 (1.07) | 1.24 (1.95) | 1.65 (2.08) | 0.75 (1.66) |
| Educational Levels | | | | | |
| No Education | 0.01 (0.12) | 0.00 (0.00) | 0.00 (0.00) | 0.00 (0.00) | 0.00 (0.00) |
| Primary | 0.10 (0.31) | 0.00 (0.00) | 0.05 (0.21) | 0.06 (0.24) | 0.03 (0.17) |
| Interm Educ | 0.40 (0.49) | 0.00 (0.00) | 0.12 (0.33) | 0.20* (0.40) | 0.03 (0.17) |
| Secondary | 0.21 (0.41) | 0.15 (0.36) | 0.22 (0.41) | 0.28 (0.45) | 0.15 (0.36) |
| Diploma | 0.06 (0.24) | 0.15 (0.36) | 0.13 (0.34) | 0.18* (0.38) | 0.07 (0.26) |
| Undergrad. | 0.13 (0.34) | 0.35 (0.48) | 0.22 (0.41) | 0.19 (0.39) | 0.25 (0.44) |
| Postgrad | 0.06 (0.24) | 0.35 (0.48) | 0.26 (0.44) | 0.09 (0.28) | 0.46 (0.50) |
| Religion | | | | | |
| Sunni Muslim | 0.76 (0.43) | 0.00 (0.00) | 0.39 (0.49) | 0.70 (0.46) | 0.01 (0.12) |
| Druze | 0.18 (0.39) | 0.82 (0.39) | 0.46 (0.50) | 0.16 (0.37) | 0.81 (0.40) |
| Other Religion | 0.06 (0.24) | 0.18 (0.39) | 0.16 (0.36) | 0.14* (0.35) | 0.18 (0.39) |
| Traits | | | | | |
| Risk Taking | 4.33 (3.61) | 3.15 (2.68) | 4.33 (3.21) | 4.60 (3.56) | 4.01* (2.71) |
| Trust dummy | 0.12 (0.33) | 0.18 (0.39) | 0.14 (0.34) | 0.16 (0.37) | 0.10 (0.31) |
| Number of Known Participants | 4.36 (6.87) | 3.83 (3.99) | 2.37 (4.36) | 2.69* (5.05) | 1.99* (3.35) |
| <i>N</i> | 67 | 66 | 147 | 80 | 67 |

Notes: Means are reported in bold and standard deviations in parantheses. Statistical differences are calculated separately for Syrians in mixed groups (1) compared to Syrians in Homogenous groups (4), and for Lebanese in mixed groups (5) compared to Lebanese in Homogenous groups (2). The difference between mixed and homogeneous groups, by nationality, is based on a regression of the outcome variable on a 'mixed' dummy. Significant differences are shown using * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Questions on trust and risk taking have been elicited through the following survey questions: 1- Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people? (Binary Options: a) Most people can be trusted. b) Need to be very careful. and 2- Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Please use the following scale, where the value 0 means: "risk averse" and the value 10 means: "fully prepared to take risks".

6. Results

In presenting our results, we first discuss participants' contribution behavior in the public goods game without and with punishment. Following this, we analyze punishment behavior in the public goods game with punishment. We finally explore individuals' welfare (as measured by their net average earnings) across treatments.

6.1 Contribution behavior in the 'No Punishment' game

Figure 1 shows the time series of average contributions across treatments for the No Punishment game. Recall that subjects first played the public good game without punishment for three rounds. We observe that in "Only Syrian" and "Mixed" treatments average contributions decline over time, while in "Only Lebanese" treatment they remain relatively stable. The observed decaying pattern of contributions as the game is repeated is in line with the vast majority of existing studies on linear standard public good games without punishment (see Chaudhuri, 2011 for an overview). Interestingly, we find that this is not the case in the "Only Lebanese" treatment suggesting that possibly these subjects have a stable notion of how much to contribute. We also see in Figure 1 and Table 2 that contribution levels are very similar comparing "Only Syrian" to "Only Lebanese" treatments, which is in contrast to previous findings that people exposed to war-related violence tend to behave more cooperatively (Voors et al. 2012, Bauer et al., 2016).

Turning to differences across treatments, we find that, in the two homogeneous treatments, average contributions are higher compared to those in the mixed groups. In particular, across all periods, average contributions in the "Only Syrians" ("Only Lebanese") treatment are equal to 5.54 (5.58) tokens, respectively. In contrast, in the "Mixed" treatment, subjects contribute on average 4.46 tokens.

Figure 1. Time series of average contributions across treatments in the no-punishment public goods game

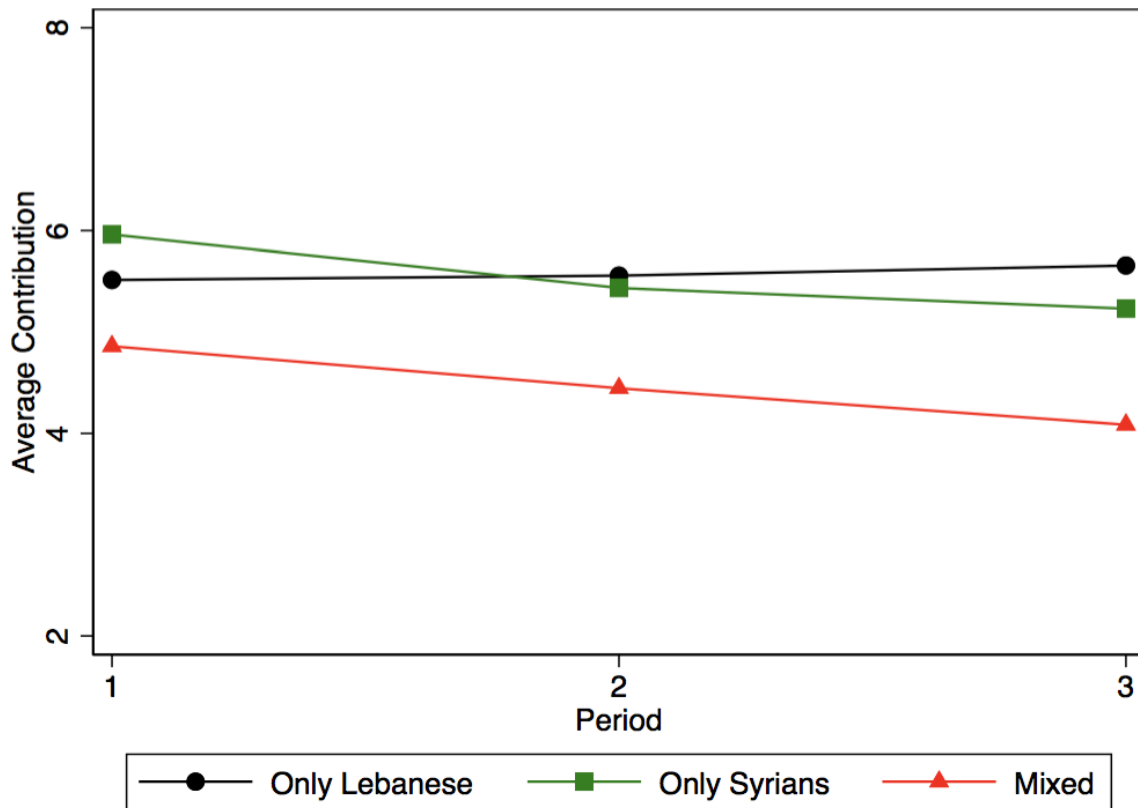


Table 2 shows the average absolute level of contribution and the average belief about others’ contribution in each treatment. By performing a Mann-Whitney test, we find that average contributions are statistically significantly different when we compare the “Only Syrians” with the “Mixed” treatments ($p = 0.006$) and the “Only Lebanese” with the “Mixed” treatments ($p = 0.002$).¹² However, we find insignificant differences in terms of contribution behavior when we compare the “Only Syrians” with the “Only Lebanese” treatments ($p = 0.935$). This indicates that homogeneous groups contribute similarly to the public good but significantly more when compared to mixed groups, thus providing support for Hypothesis 1.

We next turn to the analysis of beliefs about others’ contribution across treatments. Specifically, we observe that across all periods, average beliefs in the “Only Syrians” (“Only Lebanese”) treatment are equal to 5.72 (5.46) tokens, respectively. In contrast, in the “Mixed”

¹² All tests reported in this section are two-sided. As the unit of independent observation, we use the group level as we implemented a partners’ matching protocol in the experiment. In total, we had 156 independent observations (matching-groups), of which 39 are for the “Only Syrians” treatment, 35 for the “Only Lebanese” treatment and 82 for the “Mixed” treatment.

treatment, subjects expect others to contribute, on average, 4.84 tokens. When we test for treatment differences in beliefs about others' contribution behavior using a Mann-Whitney test, we obtain a similar pattern across treatments. In particular, we find that subjects, on average, believe that their counterpart will contribute more when we compare the "Only Syrians" with the "Mixed" treatments ($p = 0.020$) and the "Only Lebanese" with the "Mixed" treatments ($p = 0.055$). However, we find insignificant differences in terms of average beliefs about other's contribution behavior when we compare the "Only Syrians" with the "Only Lebanese" treatments ($p = 0.704$). This implies that homogeneous groups have similar beliefs about others' contribution behavior, and importantly, expect others to contribute higher amounts of tokens to the public good compared to mixed groups. Taken together, our non-parametric analysis indicates that contribution behavior and beliefs about others' contribution behavior are significantly lower in mixed groups compared to homogeneous groups.

Table 2. Average absolute levels of contribution and beliefs about contributions in the public goods game without punishment

| Treatments | Average absolute levels of contribution | Average beliefs about other's contributions |
|-------------------------|---|---|
| Only Syrians (N=39 MG) | 5.54 (1.99) | 5.72 (1.80) |
| Only Lebanese (N=35 MG) | 5.58 (1.55) | 5.46 (1.29) |
| Mixed (N=82 MG) | 4.46 (2.07) | 4.84 (1.87) |

Notes: Standard deviations in parentheses. Analysis is done at the MG (matching-group) level (total MG's N=156), the actual number of respondents is N=312.

We next perform regression analyses to check for the robustness of the treatment differences identified in our earlier non-parametric analysis. In the following, since contribution behavior and beliefs about others' contributions do not differ significantly between the "Only Syrians" with the "Only Lebanese" treatments, we pool these treatments under "Non-Mixed" treatments and compare behavior in relation to the "Mixed" treatment consisting of both Syrians and Lebanese. We test more formally for differences in contribution behavior and beliefs about others' contribution behavior across treatments using the following Tobit specification:

$$Y_i = \beta_0 + \beta_1 \text{Mixed}_i + \beta' \mathbf{X} + \gamma_t + u_i \quad (3)$$

where Y is the outcome of interest and represents either contribution made by a subject or beliefs about others' contribution, depending on the regression we report below. The reason for using a Tobit regression model is that each dependent variable is censored at 0 and 10 tokens. Our independent variables comprise: "Mixed treatment" is a dummy variable that takes 1 if the subject i is in a mixed session, and 0 if she is in a homogeneous session, period dummies which capture time patterns that may emerge from repeated play, while X is a vector of covariates that capture demographic characteristics (such as sex, age, education levels, marital status, number of children, religion, and nationality) and variables that capture self-reported characteristics of oneself – whether a participant trust others, risk taking,¹³ and the number of people they can identify by name in a given session.

Two main observations stand out from the regression analysis reported in Table 3. First, starting with column (1), we find that the coefficient of the variable "Mixed" is statistically significant at the 1% level and has a negative sign, implying that subjects in the mixed groups are less pro-social by contributing lower amounts of tokens to the public good compared to subjects in the homogeneous groups. This is in line with our observation from our non-parametric analysis. By looking separately at the Lebanese and Syrian sub-samples, as shown in columns (3) and (5), we find that the lower contribution levels are due to Lebanese subjects contributing significantly less in the mixed groups compared to the homogeneous groups. For the Syrian sub-sample, we find that the coefficient of the "Mixed treatment" has a negative sign but is not statistically significant at conventional levels. Thus, the reduction in cooperation in the mixed treatment is driven by the behavior of Lebanese participants.

¹³ Questions on trust and risk taking have been elicited through the following survey questions: 1- Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people? (Binary Options: a) Most people can be trusted. b) Need to be very careful.) and 2- Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Please use the following scale, where the value 0 means: "risk averse" and the value 10 means: "fully prepared to take risks".

Table 3. Contribution and Beliefs about other’s contributions in the Public Good Game without Punishment

| | All sample | | Lebanese sub-sample | | Syrian sub-sample | |
|-----------------|---------------------|--------------------|---------------------|-------------------|---------------------|-------------------|
| | Contribution (1) | Beliefs (2) | Contribution (3) | Beliefs (4) | Contribution (5) | Beliefs (6) |
| Mixed treatment | -1.18*** (0.37) | -0.97*** (0.31) | -1.67*** (0.46) | -1.03** (0.43) | -0.49 (0.54) | -0.98** (0.47) |
| D.Period 2 | -0.28 (0.19) | -0.06 (0.18) | 0.01 (0.28) | 0.24 (0.25) | -0.58** (0.28) | -0.34 (0.26) |
| D.Period 3 | -0.69*** (0.24) | -0.20 (0.23) | -0.17 (0.37) | 0.16 (0.34) | -1.18*** (0.31) | -0.54 (0.34) |
| Lebanese | -0.53 (0.53) | 0.14 (0.46) | | | | |
| Risk Taking | -0.05 (0.06) | -0.04 (0.05) | -0.12 (0.08) | -0.06 (0.07) | 0.02 (0.08) | -0.02 (0.06) |
| Trusting | 0.43 (0.40) | -0.14 (0.35) | 0.01 (0.51) | -0.19 (0.49) | 1.22* (0.67) | 0.21 (0.59) |
| Num known part. | 0.05 (0.04) | 0.05 (0.03) | -0.01 (0.06) | -0.04 (0.04) | 0.12** (0.06) | 0.10** (0.04) |
| Other controls | Yes | Yes | Yes | Yes | Yes | Yes |
| <i>N</i> | 840 | 840 | 402 | 402 | 438 | 438 |

Notes: Tobit estimates. Standard errors in parentheses are clustered at the matching group level. Other control variables include: sex, age, education levels, marital status, religion, and number of children. *Risk taking*: “Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?” (Scale 0-10), *Trusting*: “Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?” (Yes/No), *Num known part*: “How many participants in this experiment session do you know by name?” (Number). Treatment effects robust to different sets of control variables. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In addition, we observe the same treatment differences concerning subjects’ beliefs about other’s contribution behavior as shown in columns (2), (4) and (6). Specifically, subjects in the mixed groups expect that others will contribute less compared to subjects in the homogeneous groups. Our results provide evidence suggesting that, on average, subjects exhibit patterns of behavior in line with the concept of conditional cooperation (e.g., Fischbacher et al., 2001; Fischbacher and Gaechter, 2010), but importantly, show that conditional cooperation is

susceptible to the composition of the group. Interestingly, we find that this pattern of conditional cooperation is more salient among the Lebanese sub-sample as shown in columns (3) and (4) where the sign of the “Mixed treatment” coefficient is negative and statistically significant when the dependent variable is contribution and beliefs about other’s contribution, respectively. For the Syrian sub-sample, we observe that while they expect others to contribute less, they decrease their contributions, although the difference is not statistically significant.

Our main findings remain robust when we control for additional demographic and other control variables as explained above.

We summarize our first result below.

RESULT 1. *In the public good game without punishment, average contributions are significantly higher in the homogeneous groups compared to the mixed groups. We find the same treatment differences when we examine beliefs about other’s contributions. These patterns are mainly driven by the behavior of Lebanese participants.*

6.2 Contribution behavior in the ‘Punishment’ game

We next turn our attention to the analysis of behavior in the public goods game with punishment. Figure 2 shows the time series of average contributions in the public good game with punishment across treatments. We observe that in all three treatments average contributions remain relatively stable over time, with a slight increase for the ‘Only Syrian’ and the ‘Mixed’ treatment. The average contributions also remain lower in the “Mixed treatment” compared to the two homogeneous treatments.

Table 4 reports average contribution and beliefs about others’ contribution in the punishment game. Prior to analyzing treatment differences in the public good game with punishment, we look at whether the presence of punishment increases contributions as compared to the contribution levels in the no-punishment game, for a given treatment. Overall, we observe that, compared to the no-punishment game, average contributions significantly decrease in the presence of costly sanctioning (signtest; $p = 0.024$).¹⁴ This significant difference is driven by the observation that average contributions are significantly lower in the ‘Only Syrians’

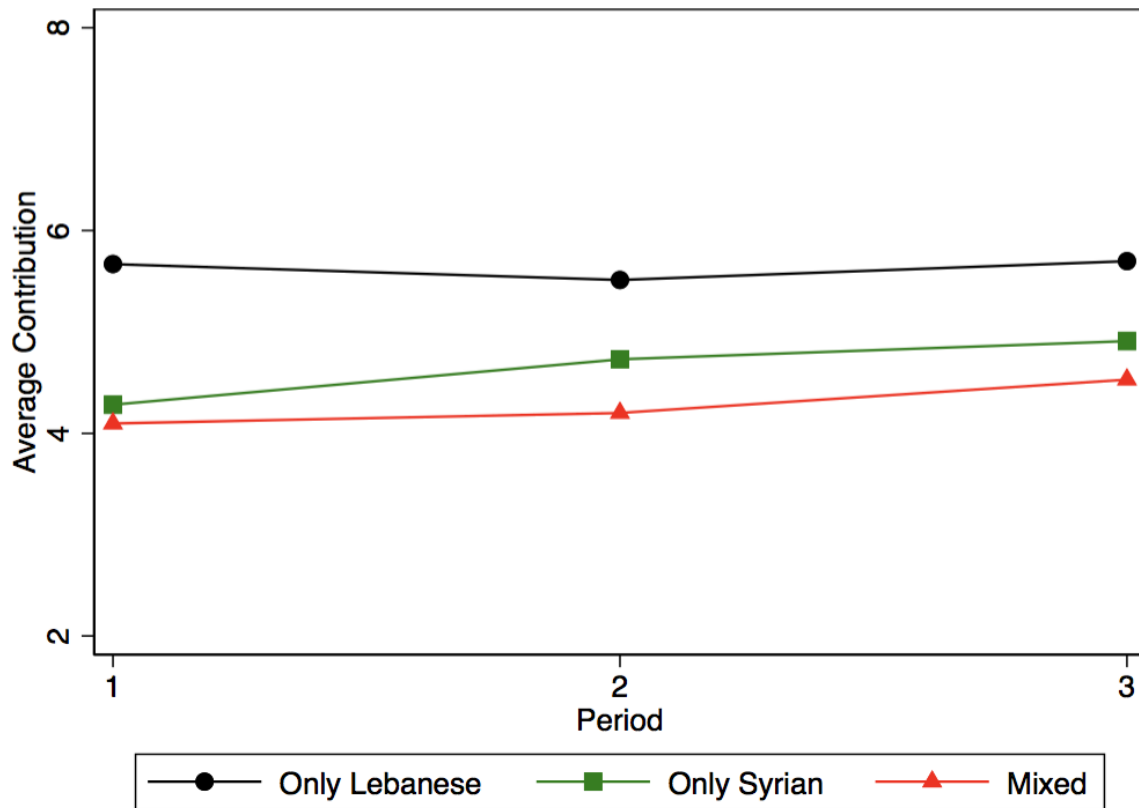
¹⁴ We also compare average contributions of period 3 of the no-punishment game, with average contributions of period 1 of the punishment game (i.e. before punishment is exercised but with the knowledge of the presence of punishment), and find that contributions are lower in the punishment game than in the no-punishment in Syrian-only sessions, not statistically different in Lebanese-only sessions, and significantly lower in mixed sessions.

treatment ($p = 0.013$). This is not the case, however, for the ‘Only Lebanese’ ($p=0.819$) and for the ‘Mixed’ treatments ($p = 0.107$), where average contributions are not significantly different between the ‘No-Punishment’ and the ‘Punishment’ treatment. This finding is consistent with previous evidence (see Herrmann et al., 2008; Gächter and Herrmann, 2011) where punishment was found to be associated with cooperation, a phenomenon referred to as “anti-social punishment”.

Our next step is to test for differences in contribution behavior across treatments. Following our observations from the no-punishment game, we examine behavior between homogeneous and mixed groups.¹⁵ On average, we find that homogeneous groups contribute 5.33 tokens and mixed groups contribute 4.28 tokens. We observe that subjects in the homogeneous groups when pooled together yield significantly higher contributions than subjects’ average contributions in the ‘Mixed’ treatment ($p = 0.013$).

¹⁵ We note that behavior in the no-punishment game is cleaner in the sense that subjects have no experience of any other game unlike behavior in the punishment game, which may be affected by the same subjects having played the no-punishment game previously.

Figure 2. Time series of average contributions across treatments in the public goods game with punishment



When we examine differences in beliefs about others' contributions across treatments, we also find that homogeneous groups when pooled together expect that others contribute significantly higher amounts to the public good than subjects' average beliefs in the 'Mixed' treatment ($p = 0.018$).¹⁶ However, our non-parametric statistical analysis does not control for potential effects that may influence contribution behavior in the 'Punishment' treatment such as demographic variables or experience stemming from playing first the no-punishment game.

¹⁶ In particular, Lebanese subjects' expectations on others' contribution in the 'Only Lebanese' treatment do not differ significantly than Syrians in the 'Only Syrians' treatment ($p = 0.505$) but are significantly more in the 'Mixed' treatment ($p = 0.013$). On the other hand, Syrians in the 'Only Syrians' treatment do not differ significantly in expectations compared to subjects in the 'Mixed' treatment ($p = 0.167$).

Table 4. Average absolute levels of contribution and beliefs about other’s contributions in the punishment public goods game

| Treatments | Average absolute levels of contribution | Average beliefs about other’s contributions |
|-------------------------|---|---|
| Only Syrians (N=39 MG) | 4.64 (2.31) | 5.18 (2.18) |
| Only Lebanese (N=35 MG) | 5.63 (2.10) | 5.50 (2.04) |
| Mixed (N=82 MG) | 4.28 (2.40) | 4.60 (2.29) |

Notes: Standard deviations in parentheses. Analysis is done at the MG (matching-group) level (total MG’s N=156), the actual number of respondents is N=312.

In Table 5, we present Tobit regression results of the contributions and beliefs in the public good game with punishment. In addition to the control variables used in the regression analysis of the no-punishment game as reported in Table 3 based on equation 3, we additionally control for differences in earnings from the no-punishment game as these may affect contribution behavior in the punishment game (recall that subjects were provided with feedback at the end of each period in the no-punishment game).

Our main finding from Table 5 is that mixed groups contribute significantly less than homogeneous groups as indicated by the coefficient of the variable “Mixed” which has a negative and statistically significant sign (see column 1). We also observe that beliefs remain lower in mixed groups compared to the homogeneous ones (as shown in column 2). These significant effects are primarily driven by the behavior of Lebanese participants (see columns 3-4). When considering the Syrian sub-sample, we observe that the coefficients of the “Mixed” variable are not significant neither for the contribution (column 5) nor for the beliefs (column 6) regressions. Finally, we find that the period dummies are generally positive and weakly significant at the 5% significance level (except for the Syrian sub-sample), suggesting that contributions weakly increase in the punishment game among Lebanese subjects.

Table 5. Contribution and Beliefs about other’s contributions in the Public Good Game with Punishment

| | All sample | | Lebanese sub-sample | | Syrian sub-sample | |
|-------------------------|---------------------|--------------------|---------------------|------------------|---------------------|-----------------|
| | Contribution (1) | Beliefs (2) | Contribution (3) | Beliefs (4) | Contribution (5) | Beliefs (6) |
| Mixed treatment | -0.99** (0.45) | -1.20*** (0.44) | -1.10* (0.61) | -1.04* (0.61) | -0.45 (0.60) | -0.80 (0.58) |
| D.Period 2 | 0.15 (0.21) | 0.38* (0.22) | 0.33 (0.34) | 0.64* (0.36) | -0.03 (0.27) | 0.12 (0.26) |
| D.Period 3 | 0.53** (0.25) | 0.58** (0.24) | 0.88** (0.42) | 0.85** (0.37) | 0.21 (0.30) | 0.33 (0.31) |
| Lebanese | 0.53 (0.58) | -0.20 (0.65) | | | | |
| Risk Taking | -0.05 (0.06) | -0.02 (0.06) | -0.09 (0.09) | 0.00 (0.09) | -0.04 (0.07) | -0.02 (0.08) |
| Trusting | 0.27 (0.52) | 0.17 (0.50) | 0.46 (0.61) | 0.57 (0.67) | 0.12 (0.91) | 0.05 (0.84) |
| Num known part. | 0.01 (0.04) | -0.06* (0.04) | 0.04 (0.05) | -0.05 (0.05) | 0.03 (0.05) | -0.04 (0.05) |
| Diff. No Pun. Earning | 0.77*** (0.07) | 0.02 (0.09) | 0.96*** (0.14) | -0.03 (0.14) | 0.71*** (0.09) | 0.06 (0.13) |
| Other control variables | Yes | Yes | Yes | Yes | Yes | Yes |
| N | 840 | 840 | 402 | 402 | 438 | 438 |

Notes: Tobit estimates. Standard errors in parentheses are clustered at the matching group level. Other control variables include: sex, age, education levels, marital status, religion, and number of children. *Risk taking*: “Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?” (Scale 0-10), *Trusting*: “Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?” (Yes/No), *Num known part*: “How many participants in this experiment session do you know by name?” (Number), *Diff. No Pun. Earning*: Average difference in earnings between the partners in the three rounds of the public good game without punishment. Treatment effects robust to different sets of control variables. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Our second result is summarized below.

RESULT 2. *In the public good game with punishment, average contributions are significantly higher in the homogeneous groups compared to the mixed groups. We find the same treatment*

differences when we examine beliefs about other's contributions. These patterns are driven by the behavior of Lebanese participants.

6.3 Punishment behavior in the 'Punishment' game

Figure 3 shows the average punishment points assigned by the punisher to another player as a function of subjects' deviation from the punisher's contribution.¹⁷ In Figure 3, each dot represents a single observation which is the average punishment across all periods at particular deviation interval as indicated in the vertical axis. Negative (positive) deviation intervals refer to cases where the punished group member's contribution is less (more) than the punisher's contribution.

A visual inspection of Figure 3 suggests that for the negative deviation interval, the punishment function has the anticipated negative slope (implying harsher punishment for larger negative deviations from the punisher's contribution behavior), as previous literature on public good games with punishment would suggest. Regarding the non-negative deviation interval, we find the substantial use of anti-social punishment. Participants are prepared to punish positive deviations from the punisher's contribution both in the homogeneous and in the mixed groups. This observation is in line with Herrmann et al. (2008) who also document the widespread punishment of co-operators in Middle East cities (e.g., Riyadh, Muscat).

We report the average social punishment (i.e. in the negative deviation interval) and anti-social punishment (i.e. in the non-negative deviation interval) in Table 6. We observe more anti-social than pro-social punishment assignment among Lebanese subjects, and vice versa among Syrian subjects. There are no significant differences between social punishment in any pairwise comparison of the three groups ("Only Lebanese" and "Only Syrian" treatments: $p=0.33$; "Only Lebanese" and "Mixed" treatments: $p=0.25$; "Only Syrian" and "Mixed" treatments: $p=0.89$). However, we observe significantly higher anti-social punishment assignment between "Only Lebanese" treatments and any of the other groups ("Only Lebanese" and "Only Syrian" treatments: $p=0.09$; "Only Lebanese" and "Mixed" treatments: $p=0.01$), with no significant difference between "Only Syrian" and "Mixed" treatments ($p=0.45$). By pooling the homogeneous groups together, we find that anti-social punishment is significantly lower in mixed groups compared to being in homogeneous groups ($p=0.04$), but this is not the case in terms of pro-social punishment ($p=0.45$). However, our non-parametric analysis does not

¹⁷ We refer to the punisher as player i and the recipient of punishment as player j .

control for the magnitude of the deviations, which is likely to be a significant determinant for assigning punishment points, as suggested by Figure 3.

Figure 3. Punishment as a function of deviation from punisher’s contribution across treatments

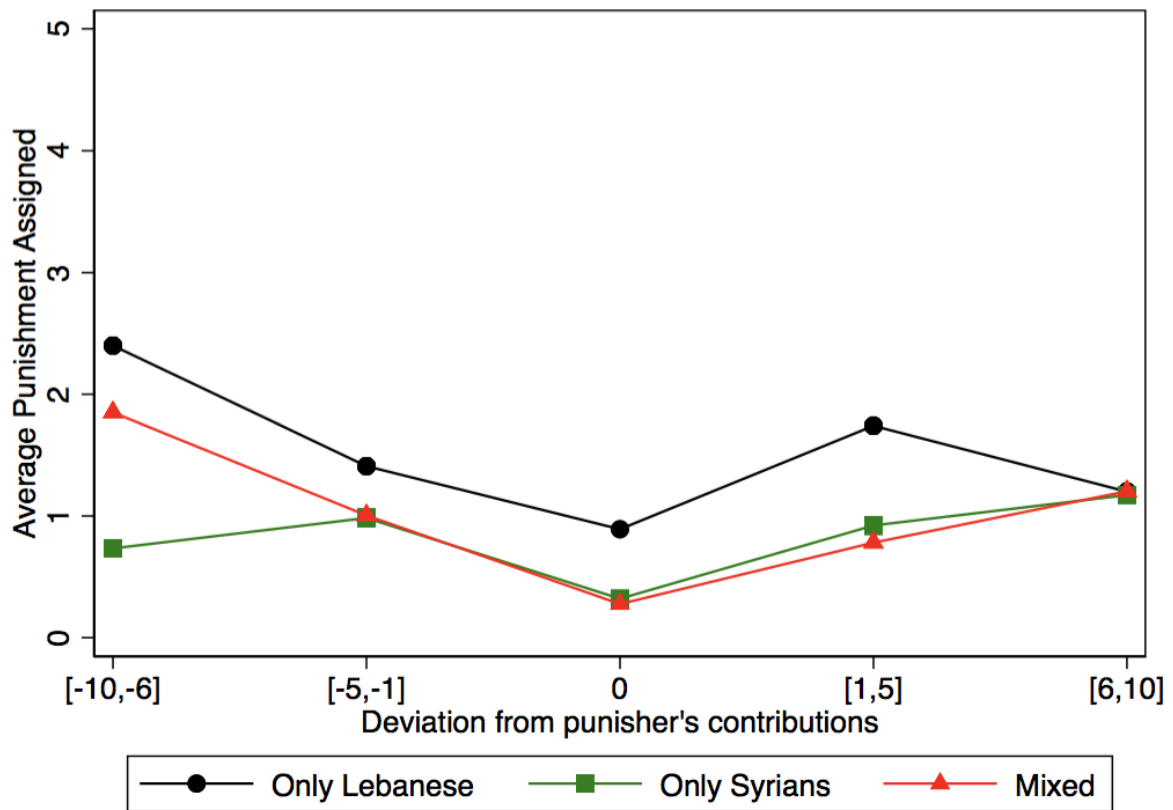


Table 6. Average punishment points assigned by a punisher to the punished group member across treatments

| | Negative Deviations | Non-negative deviations |
|-------------------------|---------------------|-------------------------|
| Only Syrians (N=39 MG) | 1.26 (1.59) | 1.38 (1.64) |
| Only Lebanese (N=35 MG) | 1.03 (1.37) | 0.85 (1.24) |
| Mixed (N=82 MG) | 1.06 (1.51) | 0.82 (1.33) |

Notes: Standard deviations in parentheses. Analysis is done at the MG (matching-group) level (total MG's N=156), the actual number of respondents is N=312.

Table 7 assesses the determinants of the assignment of punishment. To do so, we report Tobit regressions with the assignment of punishment by a subject being the dependent variable. The reason for performing a Tobit analysis is that our dependent variable, which is “punishment assigned by subject i to subject j ”, exhibits censoring at 0 and 5 points. In addition to the control variables reported in Tables 3 and 4, we include as independent variables: “Player j ’s absolute negative (contribution) deviation” and “Player j ’s positive (contribution) deviation”. Note that all deviations are calculated with respect to the punisher’s contribution. We include “absolute negative deviation” and “positive deviation” as separate regressors, since these two different sorts of deviation elicit different punishment responses as shown in Figure 3 and Table 6. The variable “absolute negative deviation” is the absolute value of the actual deviation of subject j ’s contribution from the punisher’s contribution, when subject j ’s contribution is below the punisher’s contribution; and zero otherwise. The variable “positive deviation” is constructed in an analogous way. As before, the dummy variable “Mixed treatment” equals 1 for the mixed treatment and 0 otherwise. We also included two interaction terms, which indicate whether the slope of the punishment function differs with respect to negative and positive deviations across our treatments. To gain a better understanding of the punishment patterns, we report regressions for the whole sample (columns 1 and 2) as well as the Lebanese (columns 3 and 4) and the Syrian sub-sample (columns 5 and 6), separately.

We first observe that, across most of our regression models, subjects punish significantly both negative and positive deviations from the punisher’s contribution. Specifically, the coefficient of the variable “Absolute negative deviation” has a positive sign and is statistically significant in most regressions indicating the more a subject negatively deviates from the punisher’s contribution, the harsher the punishment is. Interestingly, except for significant levels of the assignment of social punishment, we also observe that subjects engage in anti-social punishment activities. The coefficient of the variable “Positive deviation” has a positive sign and is statistically significant in most regressions reported in Table 7 (columns 1, 3, and 5). This suggests that the more a subject positively deviates from the punisher’s contribution, the harsher the punishment is. We find that the anti-social punishment behavior does not vary significantly across treatments, as shown by the interaction of the positive and negative deviations with the “Mixed treatment” (columns 2, 4, and 6). When we examine differences in how subjects use punishment across treatments, our analysis shows that subjects assign significantly less punishment in mixed groups compared to homogeneous groups as suggested by the negative and statistically coefficient of the dummy variable “Mixed treatment”.

Table 7. Punishment Assigned in the Public Good Game with Punishment

| | All sample | | Lebanese sub-sample | | Syrian sub-sample | |
|-------------------------|-------------------|-------------------|---------------------|-------------------|--------------------|--------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Mixed treatment | -0.77** (0.37) | -1.17** (0.54) | -1.09** (0.50) | -1.29* (0.77) | -0.17 (0.47) | -0.68 (0.68) |
| Abs. Negative Dev. | 0.30*** (0.07) | 0.18* (0.10) | 0.45*** (0.11) | 0.37** (0.17) | 0.18** (0.08) | 0.08 (0.11) |
| Positive Dev. | 0.27*** (0.06) | 0.25*** (0.09) | 0.43*** (0.11) | 0.43*** (0.15) | 0.18** (0.07) | 0.13 (0.11) |
| Mixed × Abs Neg Dev | | 0.21 (0.13) | | 0.14 (0.21) | | 0.20 (0.15) |
| Mixed × Pos Dev | | 0.06 (0.12) | | 0.00 (0.19) | | 0.11 (0.14) |
| D.Period 2 | -0.14 (0.19) | -0.14 (0.19) | -0.02 (0.33) | -0.03 (0.33) | -0.21 (0.21) | -0.21 (0.22) |
| D.Period 3 | -0.06 (0.24) | -0.03 (0.24) | 0.02 (0.41) | 0.03 (0.42) | -0.09 (0.27) | -0.06 (0.27) |
| Lebanese | 0.91 (0.61) | 0.93 (0.61) | | | | |
| Risk Taking | 0.03 (0.05) | 0.02 (0.05) | 0.06 (0.08) | 0.06 (0.08) | 0.02 (0.07) | 0.02 (0.07) |
| Trusting | 0.39 (0.49) | 0.44 (0.49) | 0.80 (0.67) | 0.82 (0.68) | -0.23 (0.63) | -0.14 (0.64) |
| Num known part. | -0.05 (0.03) | -0.05 (0.03) | -0.00 (0.07) | -0.00 (0.07) | -0.10*** (0.03) | -0.10*** (0.03) |
| Diff. No Pun. Earning | 0.07 (0.08) | 0.08 (0.08) | 0.08 (0.14) | 0.07 (0.14) | 0.10 (0.09) | 0.11 (0.09) |
| Other control variables | Yes | Yes | Yes | Yes | Yes | Yes |
| <i>N</i> | 840 | 840 | 402 | 402 | 438 | 438 |

Notes: Tobit estimates. Standard errors in parentheses are clustered at the matching group level. Other control variables include: sex, age, education levels, marital status, religion, and number of children. *Risk taking*: "Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?" (Scale 0-10), *Trusting*: "Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?" (Yes/No), *Num known part*: "How many participants in this experiment session do you know by name?" (Number), *Diff. No Pun. Earning*: Average difference in earnings between the partners in the three rounds of the public good game without punishment, *Abs Neg Dev*: Absolute negative deviation of partners from own contribution, *Positive Dev*: Positive deviation of partner from own contribution. Treatment effects robust to different sets of control variables. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Taken together with our Result 2, where we observed that mixed groups contribute significantly less than homogeneous groups, our finding that the mixed groups punish significantly less than the non-mixed groups suggests that the norm of sustaining high contributions is less important for mixed groups. In other words, in the mixed treatment, we observe that subjects indicate less willingness to deviate from the dominant selfish strategy compared to homogeneous groups. Overall, this suggests that norm compliance and norm enforcement is more difficult to occur in mixed groups.

Our third result is summarized below.

RESULT 3. *In the public good game with punishment, average punishment assigned is significantly lower in mixed groups compared to the homogeneous groups. We also observe the use of antisocial punishment across treatments.*

6.5 Earnings in the public good game without and with punishment

Our main findings from the previous sections show significant differences in how subjects contribute and assign punishment points across treatments. In our final section, we look at the implications that these differences have in subjects' welfare as measured by their average net earnings in each of the two games we considered. We start our analysis by examining how earnings were determined across treatments in the public good game without and with punishment in Table 8. We find that, on average, homogeneous groups have higher total earnings in the no-punishment game. However, this is not the case for the public good game with punishment where total earnings are reduced by higher levels of punishment assignment in the homogeneous groups.

Table 8. Breakdown of earnings by treatment

| | No Punishment | | Punishment | |
|-----------------------------|--------------------|-----------------|--------------------|-----------------|
| | Homogeneous groups | Mixed groups | Homogeneous groups | Mixed groups |
| Total earnings | 12.73 (2.12) | 12.23 (2.12) | 8.46 (5.80) | 8.98 (4.90) |
| Earnings after contribution | – | – | 12.50 (2.22) | 12.13 (2.27) |
| Punishment costs | – | – | -4.05 (5.12) | -3.15 (4.28) |

Notes: Earnings are measured in ECUs. Standard deviations in parentheses.

Columns 1-3 of Table 9 report three ordinary least square regression models for the no-punishment game where the dependent variable corresponds to a subject's total earnings in the public good game without punishment. Columns 4-6 report the results on total earnings of both stages of the punishment game (i.e. after contribution decisions have been made and punishment assigned), while columns 7-9 report subjects' earnings from the first (contribution) stage of the punishment game, before any punishment has been assigned in a given period. The construction of the independent variables is analogous to those we have included in Table 3 and Table 5.

Table 9. Earnings in the no-punishment and punishment games.

| | Earnings in No-Punishment Game | | | Total Earnings in Punishment Game | | | Earnings from Stage One of Punishment Game | | |
|-----------------------|--------------------------------|---------------------|-------------------|-----------------------------------|---------------------|-------------------|--|---------------------|-------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| | All sample | Lebanese Sub-Sample | Syrian Sub-Sample | All sample | Lebanese Sub-Sample | Syrian Sub-Sample | All sample | Lebanese Sub-Sample | Syrian Sub-Sample |
| Mixed treatment | -0.48*** (0.17) | -0.23* (0.23) | -0.45* (0.26) | 0.46 (0.71) | 0.85 (1.17) | -0.21 (0.82) | -0.52*** (0.19) | -0.61** (0.29) | -0.34 (0.27) |
| D.Period 2 | -0.18** (0.09) | -0.14 (0.18) | -0.22 (0.14) | 0.19 (0.33) | -0.13 (0.51) | 0.49 (0.42) | 0.05 (0.08) | -0.04 (0.15) | 0.14 (0.16) |
| D.Period 3 | -0.29*** (0.11) | -0.23 (0.19) | -0.34* (0.18) | -0.26 (0.45) | -0.37 (0.75) | -0.15 (0.49) | 0.17* (0.10) | 0.12 (0.16) | 0.23 (0.16) |
| Lebanese | 0.15 (0.33) | | | -0.38 (0.97) | | | -0.16 (0.35) | | |
| Risk Taking | -0.01 (0.03) | -0.03 (0.05) | -0.01 (0.04) | 0.05 (0.09) | 0.08 (0.15) | 0.04 (0.12) | 0.01 (0.03) | -0.03 (0.05) | 0.03 (0.05) |
| Trusting | 0.36 (0.27) | 0.20 (0.34) | 0.33 (0.43) | 0.19 (0.76) | -0.69 (1.38) | 0.92 (0.77) | 0.10 (0.25) | -0.01 (0.34) | 0.27 (0.39) |
| Num known part. | -0.01 (0.01) | 0.00 (0.02) | -0.01 (0.02) | -0.04 (0.05) | -0.07 (0.10) | 0.01 (0.06) | -0.02 (0.02) | 0.03 (0.02) | -0.03 (0.02) |
| Diff. No Pun. Earning | | | | 0.05 (0.08) | -0.24 (0.17) | 0.16 (0.10) | 0.14** (0.05) | 0.01 (0.08) | 0.19*** (0.07) |
| Other Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| N | 840 | 402 | 438 | 840 | 402 | 438 | 840 | 402 | 438 |

Notes: Least squares estimates. Standard errors in parentheses are clustered at the matching group level. Other control variables include: sex, age, education levels, marital status, religion, and number of children. Risk taking: "Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?" (Scale 0-10), Trusting: "Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?" (Yes/No), Num known part: "How many participants in this experiment session do you know by name?" (Number). Treatment effects robust to different sets of control variables. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

We find that, when considering the whole sample, earnings are significantly lower in the mixed groups, both in the no-punishment game, and in the contribution stage of the punishment game. However, by splitting the sample by nationality, we find that the reduction in earnings is mainly concentrated among Lebanese participants – who have, on average, contributed less in “Mixed” sessions (as shown in Tables 5 and 7). This reduction in average earnings can be attributed to lower contributions in the public good game. However, in the public good game with punishment (columns 4-6), we observe that there are no significant differences in total earnings between the mixed and the homogeneous treatments. While mixed groups contribute significantly less than non-mixed (leading to lower earnings from the contribution stage as shown in columns 7-9), we also find that mixed groups assign significantly less punishment than homogeneous groups. Overall, this suggests that average earnings (both from the contribution and punishment stage) among mixed and homogeneous groups do not differ significantly.

Our fourth result is summarized below.

RESULT 4. *In the public good game without punishment, average earnings are significantly lower in mixed groups compared to homogeneous groups. This is also the case in the public good game with punishment, when we only consider average earnings from the contribution stage. These patterns are driven by the behavior of Lebanese participants. If we consider total earnings including costs of punishment, there are no differences between homogeneous and mixed groups.*

6.6 Discussion

Our results are further supported by qualitative evidence based on survey data that we collected. In particular, at the end of the experiment we administered a short questionnaire to participants covering questions related to individual characteristics and attitudes toward the other group, among others. The self-reported measures of cooperation suggest that Lebanese are significantly less likely to cooperate with Syrians than Syrians are to cooperate with Lebanese.¹⁸ Thus, the self-reported measures of intergroup cooperation align with the behavior we observe in the experiment.

¹⁸ In the survey, the question on cooperation (“How likely are you to cooperate with Syrians (Lebanese)?” for Lebanese (Syrian) respondents) is measured on a three-point scale (likely – unlikely). The χ^2 test indicates a statistically significant correlation between nationality and this measure of cooperation at 1% significance level.

Furthermore, according to a national survey (UNDP, 2018), only 25% of respondents described their relations with the other group as positive or very positive in 2017 – which is when our experiment took place. In the Aley district, which is where we carried out our experiment, the report reveals that 24.2% of Lebanese and 9.5% of Syrians disagree that the two nationalities are able to work together to solve problems they have in their community. In addition, the perception of Lebanese people is such that 65% of the respondents believe that Syrians have contributed to increased crime and violence in this district, and 85% agree that the Syrians are placing strain and pressure on their public resources. There is also a clear dichotomy in the two groups' assessment of the quality of relations between the two. Indeed, around 80% of Lebanese people think that the relations are not positive, while 48% of Syrians think the same. These survey-based attitudes reveal a substantial lack of trust toward Syrian refugees, which might explain why Lebanese participants in the public good game are less cooperative in mixed groups.

Thus, overall the survey evidence collected from our sample and from a nationally representative sample is aligned with our experimental findings and provides us with further reassurance about the internal and external validity of the study.

7. Conclusion

There has been a recent interest in the impact of hosting refugees, after the Syrian conflict and the displacement of million Syrians into neighboring countries and further afield. Yet, little is known about the preferences for intergroup cooperation of refugees and their hosting communities. This paper examines cooperation and reciprocity of Syrian refugees and the Lebanese hosting community using a lab-in-the-field experimental methodology to elicit such behavior. Cooperation across groups is important for the provision of public goods and for having a stable and harmonious society. Furthermore, our context is interesting given the historical turbulent relationship between the two countries.

The results show that on average contributions are higher in homogeneous groups than in mixed groups. These results are driven mainly by the hosts (Lebanese participants), who exhibit a stronger tendency to reduce contributions in mixed groups. In terms of punishment, we find a substantial degree of antisocial punishment, especially for Lebanese, who tend to punish significantly less in mixed groups compared to in homogeneous groups. These findings

suggest that for the Lebanese hosts, although as expected there is greater cooperation and reciprocity toward own group, there is also less willingness to punish out-group defectors.

Overall, our findings indicate that mixing of the two groups leads to lower contributions to the public good, making individuals worse-off, and also that sanctions are not able to redress this lack of cooperation. Importantly, behavior is not symmetric across the two groups, as it is the host community that shows less reciprocity and cooperation toward the refugees. This highlights the challenge in many countries hosting a large number of refugees, where the host population feels anxious about competition in the labor market, and congestion in public services, and even in some cases, lack of support and as a result might behave in anti-social manner towards refugees. Hence, it is important not only to support refugees in their protracted displacement but also to provide adequate help to the host communities to alleviate any economic and social pressures. Finally, interventions aimed at increasing intergroup trust and co-operation would help to reduce own group biases and increase public good provision, which are important for the well-being of refugees and their hosting communities.

REFERENCES

- Akerlof, G.A. and Kranton, R.E., 2000. Economics and identity. *Quarterly Journal of Economics*, 115(3): 715-753.
- Alesina, A. and Ferrara, E.L., 2005. Ethnic diversity and economic performance. *Journal of Economic Literature*, 43(3): 762-800.
- Alexander, M. and Christia, F., 2011. Context modularity of human altruism. *Science*, 334(6061):1392-1394.
- Alix-Garcia, J., Walker, S., Bartlett, A., Onder, H. and Sanghi, A. 2018. Do refugee camps help or hurt hosts? The case of Kakuma, Kenya. *Journal of Development Economics*, 130 : 66–83.
- Assaad, R., Ginn, T. and Saleh, M. 2019. Impact of Syrian Refugees on Education Outcomes in Jordan, *CEPR DP* 14056.
- Balkan, B. and Tumen, S. 2016. Immigration and Prices: Quasi-Experimental Evidence from Syrian Refugees in Turkey, *Journal of Population Economics*, 29(3): 657–686.
- Balkan, B., Tok, E. O., Torun, H. and Tumen, S. 2018. Immigration, Housing Rents, and Residential Segregation: Evidence from Syrian Refugees in Turkey, *IZA DP* No. 11611.
- Bartos, V. and Lively, I. 2018. Sanctioning and trustworthiness across ethnic groups: Experimental evidence from Afghanistan, *CESifo Working Paper Series* No. 7179.
- Bauer, M., Blattman, C., Chytilová, J., Henrich, J., Miguel, E. and Mitts, T., 2016. Can War Foster Cooperation? *Journal of Economic Perspectives*, 30(3): 249–274.
- Becker, S.O. and Ferrara, A., 2019. Consequences of forced migration: A survey of recent findings. *Labour Economics*, 59: 1-16.
- Bernhard, H., Fischbacher, U., Fehr, E., 2006. Parochial altruism in humans. *Nature*, 442: 912-915.
- Borjas, G. J. 2017. The Wage Impact of the Marielitos: A Reappraisal. *Industrial and Labor Relations Review*, 70(5): 1077–110.
- Card, D. 1990. The Impact of the Mariel Boatlift on the Miami Labor Market. *Industrial and Labor Relations Review* 43 (2): 245-57.

- Castro, M.F., 2008. Where are you from? Cultural differences in public good experiments. *The Journal of Socio-Economics*, 37(6): 2319-2329.
- Cettolin, E. and Suetens, S. (forthcoming) Return on Trust is Lower for Immigrants. *Economic Journal*.
- Chakravarty, S. and Fonseca, M.A., 2014. The effect of social fragmentation on public good provision: an experimental study. *Journal of Behavioral and Experimental Economics*, 53:1-9.
- Chaudhuri, A., 2011. Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics*, 14: 47–83.
- Clemens, M. and Hunt, J. 2019. The labor market effects of refugee waves: reconciling conflicting results, *Industrial Labor Relation Review*, 72(4): 818-857.
- Cox, J.C. and Orman, H. W., 2015. Trust and Trustworthiness of Immigrants and Native-born Americans. *Journal of Behavioral and Experimental Economics*, 57: 1-8.
- Del Carpio, X. V., and Wagner, M. 2015. The Impact of Syrian Refugees on the Turkish Economy: Regional Labour Market Effects. *World Bank Policy Research Paper No. 7402*. Washington, DC.
- Eckel, C.C. and Grossman, P.J., 2005. Managing diversity by creating team identity. *Journal of Economic Behavior & Organization*, 58(3): 371-392.
- Fallah, B., Krafft, C. and Wahba, J. 2019. The Impact of Refugees on Employment and Wages in Jordan. *Journal of Development Economics*, 139: 203-216.
- Fehr, E., Gächter, S., 2000. Cooperation and punishment in public goods experiments. *American Economic Review*, 90 (4): 980-994.
- Fehr, E., Gächter, S., 2002. Altruistic punishment in humans. *Nature*, 415 (6868): 137-140.
- Fehr, E. and Schmidt, K.M., 2006. The economics of fairness, reciprocity and altruism—experimental evidence and new theories. *Handbook of the economics of giving, altruism and reciprocity*, 1, 615-691.
- Felfe, C., Kocher, M.G., Rainer, H., Saurer, J. and Siedler, T., 2018. More opportunity, more cooperation? The behavioral effects of birthright citizenship on immigrant youth. CESifo Working Paper Series No. 6991.

- Fershtman, C. and Gneezy, U., 2001. Discrimination in a segmented society: An experimental approach. *The Quarterly Journal of Economics*, (1): 351-377.
- Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics letters*, 71(3), 397-404.
- Fischbacher, U., & Gächter, S. (2010). Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *American economic review*, 100(1), 541-56.
- Gächter, S., Renner, E. and Sefton, M., 2008. The long-run benefits of punishment. *Science*, 322(5907):1510-1510.
- Gächter, S., and B. Herrmann 2011. The limits of self-governance when cooperators get punished: Experimental evidence from urban and rural Russia. *European Economic Review*. 55(2): 193-210.
- Gneezy, U. and Imas, A. (2016). 'Lab in the Field: Measuring Preferences in the Wild', in (A. Banerjee and E. Duflo, eds.), *Handbook of Field Experiments*, Elsevier.
- Goette, L., Huffman, D. and Meier, S., 2006. The impact of group membership on cooperation and norm enforcement: Evidence using random assignment to real social groups. *American Economic Review*, 96(2): 212-216.
- Goette, L., Huffman, D. and Meier, S., 2012. The impact of social ties on group interactions: Evidence from minimal groups and randomly assigned real groups. *American Economic Journal: Microeconomics*, 4(1):101-15.
- Habyarimana, J., Humphreys, M., Posner, D.N. and Weinstein, J.M., 2007. Why does ethnic diversity undermine public goods provision?. *American Political Science Review*, 101(4): 709-725.
- Herrmann, B., Thöni, C. and Gächter, S., 2008. Antisocial Punishment Across Societies. *Science*, 1362-1367.
- Malaeb, B. (2018). "State Fragility in Lebanon: Proximate Causes and Sources of Resilience". LSE-Oxford Commission on State Fragility, Growth and Development Report.
- Maystadt, J-F, and Verwimp, Philip. 2014. Winners and Losers among a Refugee-Hosting Population. *Economic Development and Cultural Change*, 62 (4): 769-809.
- Miguel, E. and Gugerty, M.K., 2005. Ethnic diversity, social sanctions, and public goods in Kenya. *Journal of Public Economics*, 89(11-12): 2325-2368.

- Ruffle, B.J. and Sosis, R., 2006. Cooperation and the in-group-out-group bias: A field test on Israeli kibbutz members and city residents. *Journal of Economic Behavior & Organization*, 60(2): 147-163.
- Ruiz, I. and Vargas-Silva, C., 2015. The labor market impacts of forced migration. *American Economic Review Paper Proceedings*, 105 (5) : 581-586.
- Tumen, S., 2016. The Economic Impact of Syrian Refugees on Host Countries: Quasi-Experimental Evidence from Turkey. *American Economic Review* 106 (5): 456–460.
- Tumen, S., 2019. Refugees and ‘Native Flight’ from Public to Private Schools, *Economics Letters*. 181: 154–159.
- Tajfel, H., and Turner, J.C. 1979. “ An Integrative Theory of Intergroup Conflict.” In *Social Psychology of Intergroup Relations*, ed. W. G. Austin and S. Worchel. Monterey : Brooks/Cole, pp. 33– 47.
- United Nation Development Programme (UNDP) 2018. “Regular Surveys on Social Tensions Throughout Lebanon”. Report, data2.unhcr.org/en/documents/download/67048 . Accessed 20 Jul. 2019.
- United Nations Higher Commissioner for Refugees 2018. *Global Trends: Forced Displacement in 2018*, Geneva: UNHCR.
- United Nations Higher Commissioner for Refugees, 2019. UNHCR Statistics – The World in Numbers. Retrieved from <http://popstats.unhcr.org/en/overview>
- Verme, P. and Schuettler, K. 2019. The Impact of Forced Displacement on Host Communities: A Review of the Empirical Literature in Economics, *Policy Research Working Paper 8727*, Washington DC: the World Bank.
- Voors, M. J., Nillesen, E. M, Verwimp, P., Bulte, E. H., Lensink, R. and Van Soest D. P., 2012. Violent Conflict and Behavior: A Field Experiment in Burundi, *American Economic Review* 102(2): 941–964.

Cooperation in a Fragmented Society: Experimental Evidence on Syrian Refugees and Natives in Lebanon

Appendix 1 – For online publication

Table A1. Contribution and Beliefs about other’s contributions in the Public Good Game without Punishment – without any control variables (apart from period fixed effects)

| | All sample | | Lebanese sub-sample | | Syrian sub-sample | |
|-----------------|---------------------|----------------|---------------------|----------------|---------------------|----------------|
| | Contribution (1) | Beliefs (2) | Contribution (3) | Beliefs (4) | Contribution (5) | Beliefs (6) |
| Mixed treatment | -1.21*** | -0.82** | -1.69*** | -0.78* | -0.86 | -0.89* |
| | (0.37) | (0.32) | (0.44) | (0.41) | (0.54) | (0.47) |
| D.Period 2 | -0.43** | -0.24 | -0.13 | 0.09 | -0.70** | -0.54** |
| | (0.20) | (0.18) | (0.29) | (0.25) | (0.29) | (0.25) |
| D.Period 3 | -0.72*** | -0.31 | -0.27 | 0.01 | -1.11*** | -0.58* |
| | (0.23) | (0.22) | (0.37) | (0.35) | (0.30) | (0.30) |
| <i>N</i> | 936 | 936 | 420 | 420 | 516 | 516 |

Notes: Tobit estimates. Standard errors in parentheses are clustered at the matching group level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A2. Contribution and Beliefs about other’s contributions in the Public Good Game without Punishment – without any control variables (apart from period fixed effects)

| | All sample | | Lebanese sub-sample | | Syrian sub-sample | |
|-----------------|---------------------|----------------|---------------------|----------------|---------------------|----------------|
| | Contribution (1) | Beliefs (2) | Contribution (3) | Beliefs (4) | Contribution (5) | Beliefs (6) |
| Mixed treatment | -0.90* | -0.80* | -1.50** | -1.01* | -0.34 | -0.61 |
| | (0.47) | (0.43) | (0.63) | (0.58) | (0.65) | (0.55) |
| D.Period 2 | 0.14 | 0.41** | 0.30 | 0.64* | 0.01 | 0.22 |
| | (0.20) | (0.21) | (0.33) | (0.35) | (0.26) | (0.25) |
| D.Period 3 | 0.54** | 0.67*** | 0.79* | 0.87** | 0.33 | 0.50 |
| | (0.25) | (0.24) | (0.42) | (0.36) | (0.29) | (0.32) |
| <i>N</i> | 936 | 936 | 420 | 420 | 516 | 516 |

Notes: Tobit estimates. Standard errors in parentheses are clustered at the matching group level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A3. Punishment Assigned in the Public Good Game with Punishment– without any control variables (apart from period fixed effects)

| | All sample | Lebanese sub-sample | Syrian sub-sample |
|-----------------|------------|---------------------|-------------------|
| | (1) | (3) | (5) |
| Mixed treatment | -0.64* | -0.94* | -0.32 |
| | (0.37) | (0.53) | (0.48) |
| D.Period 2 | -0.18 | -0.07 | -0.27 |
| | (0.18) | (0.33) | (0.21) |
| D.Period 3 | -0.09 | 0.02 | -0.18 |
| | (0.23) | (0.41) | (0.25) |
| <i>N</i> | 936 | 420 | 516 |

Notes: Tobit estimates. Standard errors in parentheses are clustered at the matching group level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A4. Earnings in the punishment and no-punishment games- No controls.

| | Earnings in No-Punishment Game | | | Total Earnings in Punishment Game | | | Earnings from Stage One of Punishment Game | | |
|-----------------|--------------------------------|---------------------|-------------------|-----------------------------------|---------------------|-------------------|--|---------------------|-------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| | All sample | Lebanese Sub-Sample | Syrian Sub-Sample | All sample | Lebanese Sub-Sample | Syrian Sub-Sample | All sample | Lebanese Sub-Sample | Syrian Sub-Sample |
| Mixed treatment | -0.50*** | -0.58** | -0.45** | 0.53 | 0.87 | 0.17 | -0.36* | -0.79*** | -0.01 |
| | (0.15) | (0.24) | (0.22) | (0.65) | (1.01) | (0.76) | (0.18) | (0.27) | (0.27) |
| D.Period 2 | -0.16** | -0.09 | -0.21* | 0.09 | -0.15 | 0.28 | 0.06 | -0.03 | 0.13 |
| | (0.08) | (0.17) | (0.13) | (0.31) | (0.49) | (0.38) | (0.07) | (0.14) | (0.14) |
| D.Period 3 | -0.27*** | -0.20 | -0.33** | -0.14 | -0.26 | -0.05 | 0.19** | 0.12 | 0.25* |
| | (0.09) | (0.18) | (0.15) | (0.41) | (0.71) | (0.43) | (0.09) | (0.16) | (0.14) |
| <i>N</i> | 936 | 420 | 516 | 936 | 420 | 516 | 936 | 420 | 516 |

Notes: Earnings from Stage One of Punishment Game , Columns 7-9, are before subjects have assigned punishment to each other. Standard errors in parentheses are clustered at the matching group level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix 2

Figure A1. Image of the setup of the experiment room

