Overcoming the Trust Deficit:
Intergroup Contact and Associational Life in post-ISIS Iraq

Salma Mousa[*]

April 11, 2018

Abstract

Can intergroup contact create trust after conflict? This study experimentally tests the potential for intergroup contact to reduce prejudice in a conflict setting for the first time. I do this by randomly assigning Iraqi Christians displaced by ISIS either to an all-Christian soccer team or to a team mixed with Sunni Arabs – who share the same ethno-religious background as ISIS. I find persistent changes to both trusting attitudes and behaviors: Christians assigned to mixed teams are 26.1 percentage points more likely to attend a social event open to Muslims three weeks after the intervention, 45.5 percentage points more likely to train with Muslims four months later, and 0.31 standard deviations more likely to believe that peaceful coexistence is possible. These findings point to the potential for civic organizations that cut across social cleavages to rebuild social trust after ethnic conflict.

Key Words: Religion and politics, comparative politics, ethnic conflict, Middle East, democracy, human rights

In June 2014, ISIS committed mass atrocities against minorities from the Nineveh plains in what was later deemed a genocide by the United Nations and the United States. ISIS raids displaced at least 10,000 Christians to Iraqi Kurdistan within a matter of days (Al-Ameen 2014). Many Christians believe their Muslim neighbors were complicit in the raids, leading to intense distrust, a fear of returning home even to liberated areas, support for local self-defense militias, and potential for reprisal killings and future conflict (Center for the Prevention of Genocide 2016). Despite the importance of social trust to security, good governance, economic development, and democracy (Putnam 2007; Fukuyama 1995; Fehr, Kosfeld and Fischbacher 2005; Rothstein and Uslaner 2005; Knack and Keefer 1997), little is known about how to build social trust – especially after violent conflict.

Scholars have recently turned to field experiments to test the potential for intergroup contact to drive social trust (Enos 2014). Prolonged, positive, and cooperative contact reduced prejudice toward black students in South Africa (Burns, Corno and La Ferrara 2015), between Muslim and Christian students in Nigeria (Scacco and Warren 2016), and toward poor students (Rao et al. 2013) and other caste groups in India (Lowe 2017), whereas interventions imposing brief or adversarial contact worsened prejudice toward the poor (Sands 2017) and ethnic minorities (Enos 2014) in the U.S. Collectively, these studies suggest that extended, collegial contact is promising. Less is known, however, about the effect of intergroup contact in conflict settings: fewer than 3% of the 515 intergroup contact studies reviewed in Pettigrew and Tropp's (2006) meta-analysis involved groups in conflict. These same societies arguably have the most to gain from successful intergroup contact interventions, especially given the potential for conflict to spread to new regions via the experiences carried by displaced people (Ditlmann and Samii 2016; Salehyan and Gleditsch 2006). I address this gap using a field experiment to test the effects of contact between actively antagonistic groups. To do this, I randomly assign extended contact with Sunni Arab Muslims among a sample of amateur Christian soccer players displaced by ISIS – a Sunni Arab group. Unlike most contact interventions that rely on survey responses or lab-in-the-field behaviors measured on the same day as the intervention, I record both attitudes and behaviors measured up to four months after the intervention concludes to capture persistent changes in trust (Paluck, Green and Green 2017).

Despite the atrocities suffered at the hands of ISIS and an intense distrust of Muslims even when compared to other displaced Christians, I find causal evidence that positive, cooperative, and prolonged intergroup contact – conditions I label as "meaningful contact" – can bridge the trust deficit created by war. I find that Christians randomly assigned to soccer teams with Muslim players are 26 percentage points (50%) more likely to attend a mixed social event, 46 percentage points (353%) more likely to train with Muslims four months after the initial treat-

ment has ended, and 31% more likely to believe in the possibility of peaceful coexistence. Heterogenous effects further point to the potential for intergroup contact to build social trust: those with the least prior contact with Muslims and the most prejudiced baseline attitudes experienced the largest treatment effects. I present qualitative evidence demonstrating that Muslim players were able to build and transmit good reputations that disassociated them from ISIS through repeated interactions with Christian teammates over time. Moreover, the trust-building effects of contact do not appear to come with the downside of increased ethno-centrism or bouts of intergroup aggression among those on all-Christian teams. I leverage the random assignment of a Muslim vs. Christian referee to present match-level analyses and qualitative evidence demonstrating that participants deliberately prevented interethnic conflict between players and referees of different religions, therefore preserving the positive effects of civic participation on cooperativeness (Putnam, Leonardi and Nanetti 1994). This ingroup policing strategy was complemented by emerging norms against hateful speech. I build a theory of change based on these insights. I propose that overlapping interests between groups can be leveraged to unlock the trust-building potential of contact. Over time, sufficiently incentivized and meaningful contact fills information gaps about the "other," relieves out-group anxiety, and induces empathy at the individual level.

These results inform a largely observational debate on whether conflict hinders or helps social trust. Some scholars argue that wartime violence can severely undermine the potential for collective action and social trust (Wood 2003; Collier and Esteban 2007; Walter 2004; Posen 1993) while others maintain that violent conflict may strengthen resilience and social cohesion through "collective coping" mechanisms (Bellows and Miguel 2009; Rohner, Thoenig and Zilibotti 2013; Blattman 2009). I tie these theories together by demonstrating that conflict can deal a blow to social capital but that intergroup contact within the equalizing setting of cross-cutting associations can rebuild lost trust even in the absence of shared victimhood. Others still contend that prejudice has a limited role in shaping the structural factors that ultimately create ethnic conflict (Green and Seher 2003; Forbes 1997). In contrast, I show that contact can spur psychological processes that change deeply engrained beliefs and feelings about the outgroup, the ingroup, and the relationship between the two (Bar-Tal 2000). Transforming attitudes is important given that individual- and group-level prejudice can perpetuate violence even after the initial causes of conflict have become irrelevant (Deutsch 1977; Tajfel and Turner 1979). The positive effects I find as a result of randomly assigning contact may also help explain mixed outcomes documented in observational studies of interethnic peacebuilding programs both in and outside of the sports sphere (Hammack 2006; Schroeder and Risen 2016; Leitner, Galily and

2

Shimon 2012; Hewstone et al. 2006; Varshney 2003).[1] More broadly, I present causal evidence regarding civic life and social capital from an Arab context where authoritarian persistence is often attributed to the weakness of associational life (Bellin 2012; Jamal 2009).

Overcoming the Trust Deficit: a Theory of Change

I define social trust as follows: an equilibrium in which everyday interactions between two groups are not conditioned on identity. The goal is thus to reduce the role of identity in preventing intergroup interactions. While structural factors such as residential segregation and political discourse also shape social trust, I focus on intergroup cooperation as an in-road to building social cohesion at the grassroots level. The theory presented here can be applied to any two antagonistic groups, be it ethnic groups who have co-existed for centuries, or relatively new arrivals in Western societies. Within a social identity group, we have these three stylized types: (1) "never integrators," who never interact with the out-group even if necessary, (2) "always integrators," who always interact with the out-group even if unnecessary, and (3) "potential integrators," who interact with the out-group only when necessary. Members of this group will deal with the out-group when needed, as customers, government employees, or service workers, for instance, but little more. Potential integrators are most amenable to trust-building interventions, given their low – but not prohibitively low – baseline levels of trust. With the right incentives, potential integrators can be encouraged to interact with the outgroup. Tapping into an interest shared by potential integrators and the out-group is a first step toward setting up such an incentive structure.

I tie insights from three literatures: (1) that positive, prolonged, and cooperative intergroup contact – what I term "meaningful contact" can reduce prejudice, (2) that cross-cutting associations can create "bridging" trust between groups (Varshney 2003; Putnam, Leonardi and Nanetti 1994), and (3) that a clear economic stake in cooperation can promote pro-peace attitudes (Jha and Shayo 2017). I start with the premise that group boundaries are flexible (Brubaker and Cooper 2000). Locally relevant incentives can hook potential integrators from antagonistic groups into cooperating in mutually beneficial endeavor – like business associations, reading clubs, sports teams, or trade unions. Meaningful contact within these settings unlocks the trust-building potential of social contact in two ways. First, meaningful contact fills information gaps about the "other," relieves out-group anxiety, and induces

---

[1] Two inter-ethnic sports programs in Israel/Palestine led to soccer coaches abandoning a peace education and training program (Litvak-Hirsch, Galily and Leitner 2016) and worsened outgroup affect among Palestinian basketball playing youth while improving it among their Jewish peers (Ditlmann and Samii 2016).

empathy. Potential integrators will, on average, begin to humanize their out-group partners and misperceptions and will decrease. Second, because the endeavor itself is based on a shared interest, potential integrators will begin to recognize a source of common ground. Not all shared identities are salient, but a joint stake in investing in this overlapping identity helps make it so. Without changing the salience of existing identities, meaningful contact can thus activate a new, bridging identity while imposing intergroup exposure under ideal conditions for prejudice reduction. The outcome of this process is increased willingness to interact with the out-group – at first, only in the realm of mutual benefit based on a shared interest, but later, perhaps a widened range of activities. This process thus leads to cooperation in the short-term, and more tolerant and trusting attitudes and behaviors toward the out-group (intervention participants and the out-group at large) in the mid- to long-term.

What must be true for in order for this process to build lasting trust? Four key scope conditions must be met. First, there cannot be an imminent threat of violence as a result of interacting with the out-group. Second, the out-group must be geographically accessible – the transaction costs of interacting should not be prohibitive. Third, the endeavor must meet the "mutually beneficial" criteria. A rule of thumb here is that, absent identity, participation in this endeavor would occur. Given that negative or adversarial contact worsens trust, (Lowe 2017; Allport 1979; Enos 2014; Sands 2017), the endeavor should not involve competition between antagonistic groups (Jha 2013). Four, related to Allport's 1979 criterion of approval from communal authorities, both communities should deem participation in this endeavour to be minimally acceptable: there should be no prohibitive stigma against participation. An endeavor that is clearly beneficial to in-group members, such as cooperation in a trade union, can weaken taboos against intergroup contact in this regard. Absent these barriers, potential integrators can theoretically be incentivized to engage in meaningful contact. Whether or not leveraging an existing interest (e.g. soccer) as opposed to creating it (e.g. assigning groups to the same stock portfolio (Jha and Shayo 2017) or rotating credit associations (Karlan 2007)) is more conducive to building trust represents a fruitful avenue for further research.

Hypotheses and Observable Implications

The theory predicts that meaningful contact that leverages a shared interest, coupled with the right incentives to encourage committed participation, can reduce psychological barriers to social trust. This experiment tests two of features of meaningful contact – the level (Muslim teammates vs. no Muslim teammates) and nature (cooperative vs. adversarial) of intergroup contact. The treatment is bundled, as being assigned to the treatment condition entails the incorporation of Muslim players as well as imposing the need to cooperate with these players.

4

- *H1*: Previous level of contact with Muslims moderates treatment effects.

- *H2a*: Given the negative impact of adversarial contact, no over-time changes to trusting attitudes should be observed among control participants.

- *H2b*: Given the positive impact of secondary contact, positive over-time changes to trusting attitudes should be observed among control participants.

- *H3*: Membership in other civic organizations amplifies treatment effects.

- *H4*: Those with the most positive assessment of the league experience overall should be associated with higher treatment effects.

The centrality of intergroup contact to this theory gives rise to the first of four hypotheses. One could expect heterogenous effects based on the level of contact with Muslims prior to the experiment (*H1*). Ex-ante predictions about the direction and magnitude of this effect are unclear, however: those with the least prior contact with Muslims potentially have the most to gain from a mixed team environment, but attitudes may be sticky if low contact is itself driven by a high level of prejudice. Second, the contact theory stipulates that a common aim is needed to unlock the prejudice-reducing potential of contact (Allport 1979; Pettigrew 2016). Another observable implication is therefore that adversarial contact with Muslims – in this case, control teams facing off against mixed teams with Muslims in their ranks – will reduce trust (*H2a*). I test this hypothesis by comparing T1 and T2 attitudes for control participants as well as comparing the prevalence of red and yellow cards in matches between two control teams relative to games between one control and one mixed team. If mutual cooperation is a necessary condition for positive intergroup contact, then one would observe nil or negative changes among control participants over time, and heightened match-level aggression when control teams are pitted against Muslim players.

A third hypothesis predicts precisely the opposite effects of the adversarial contact theory. Scholars have shown that secondary (Cameron et al. 2006), vicarious (Simonivitz, Kezdi and Kardos 2017), and even imagined (Crisp and Turner 2009) contact with outgroup members can encourage perspective-taking and reduce prejudice. The secondary contact hypothesis predicts that exposure to Muslims as a result of competing in the same league can build trust (*H2b*). If the secondary contact hypothesis is true, three observable results follow: positive attitudinal changes from T1 to T2 for control participants, no changes in match-level aggression when control teams play against mixed teams, and some control teams choosing to train with Muslim players four months after the intervention concludes. Making an effort to connect with Muslim players outside of one's team and arrange practice at least once a week is a particularly costly outcome for control participants relative to their treated counterparts.

5

Moreover, if secondary contact can build trust, then participating in a mixed global environment such a tournament would seem to confer some of the trust-building benefits associated with cross-cutting groups. Positive shifts in trusting attitudes for control participants would point to the potential for low doses of intergroup contact to build social trust, albeit to a weaker extent compared to higher dose treatments like playing on the same team. Fourth, the social trust literature characterizes civic cultures as re-enforcing. A related prediction is thus that those who report membership in another associations, especially cross-cutting ones, are more likely to experience stronger treatment effects (*H3*). Lastly, negative contact has been found to disproportionately shape tolerance relative to positive contact (Graf, Paolini and Rubin 2014). A final hypothesis is that those who were dissatisfied with their experience in the league are less likely to experience strong treatment effects (*H4*).

ISIS Genocide and Post-Displacement Context

Iraq has been home to an indigenous Christian community for almost two millennia, making them one of the oldest continuous Christian communities in the world. Most Iraqi Christians consider themselves ethnic Assyrians, speak both Eastern Aramaic and Arabic, and subscribe to the Chaldean Catholic Church. Iraq's Christians are thus ethnically, religiously, and linguistically distinct from their compatriots. Nevertheless, Christians have traditionally played a prominent role in Iraqi public life – counting the founder of the Ba'ath Party and Saddam Hussein's Foreign Minister, Tareq Aziz, among their ranks. This same association with the Ba'ath regime fuelled the perception of Christians as collaborators and led to their targeting when the Hussein regime was toppled in 2003. Christians subsequently poured out of Iraq, dwindling from around 1.5 million in 2003 (7% of the population) to as low as 300,000 before the arrival of ISIS in 2013 (Basu 2016).

The summer of 2014 marked the scale-up of ISIS' ethnic cleansing campaign against Nineveh's minorities, leading to the systematic enslavement and slaughter of half a million Yazidis, Christians, and Shi'ites, and the forcible displacement of other groups such as Arabs, Kurds and Turkmen of both Shi'ite and Sunni backgrounds. Those expelled by ISIS – Muslim and Christian – joined the three million Iraqis internally displaced by heavy fighting between the Iraqi Security Forces and armed groups from December 2013 to April 2017 (UNHCR 2017). Christian strongholds in the Nineveh governorate, formerly the most diverse province in Iraq, were under ISIS control until October 19, 2016 when they were liberated as part of the Battle of Mosul. Many of the displaced fled to the Christian suburb of Ankawa in the Kurdish city of Erbil, roughly 50 miles away from Mosul. Displaced Muslims

and Christians in Ankawa reside either in camps run by faith organizations in cooperation with the UNHCR or in private settlements such as living with host families or in rented accommodation.[2] Camps and residential neighborhoods are segregated by religion with little opportunity for intergroup contact – security checkpoints often prevent Muslims from accessing Christian camp areas, and vice versa. Both the Christian distrust of Muslims and the broader trend of latent Muslim intolerance toward Christians (Lugo et al. 2013) imply that a common displacement experience seems to have done little to ease tensions in Erbil.

Experiment Design and Measurement

The experiment took place between March 27 and June 2, 2017 in Ankawa, a Christian suburb of Erbil that hosts displaced Iraqis of all backgrounds. Like much of life in Erbil, the city's 30 amateur male soccer teams are segregated by religion. I randomly chose fourteen teams – founded by displaced Christians – to participate in the league. Captains were told that a well-known local NGO was setting up a ten-week soccer league for displaced people in the area, with two conditions for participating.[3] First, each team had ten players but were told that fourteen are needed to join the league, and that the additional four players may or may not be Christian. Second, all players agree to complete a brief survey on the displacement experience and their views on Iraqi society before and after the league. All contacted teams accepted despite initially protesting the inclusion of Muslims.[4] Treated teams thus received an additional four Muslim players each while control teams instead received four fellow Christians. Incentives designed with the team captains, including professional referees, reserved fields, and trophies awarded to the top three teams, led to perfect compliance and committed participation throughout the ten weeks (Figure 1).[5] The vast majority (86.4%) of contacted participants were retained until the end of the study while the remaining 14.6% dropped out before treatment assignments were made.[6]

I randomly drew the four players added to each team from rosters of amateur teams not chosen to partake in the study, blocking on age range and region. Recruiting Muslims and Christians drawn from other amateur teams

---

[2]Of the 1,163 displaced Christians in the Erbil area surveyed by the author between March and August 2017, around 13% lived in camps.

[3]The Humanitarian Nineveh Relief Organization, an Iraqi Christian-led NGO serving IDPs and the operational partner for this study.

[4]See "Results and Discussion" for a deeper elaboration on initial hostility toward Muslims.

[5]Only 2% of matches played were forfeited due to absenteeism.

[6]Interviews with the 14.6% of initially contacted players who were forced to drop out before treatment assignments were made revealed that the Baghdad government compelled them to return to Nineveh once their houses were deemed to be sufficiently reconstructed. By definition, pre-treatment characteristics do not correlate with treatment assignment in this case. The study thus has full compliance (i.e. no drop outs post-assignment) from the perspective of randomization, or random missingness.

Figure 1: Photo of a typical league game, March 27, 2017

Matches were attended by hundreds of spectators, further enhancing players' stake in cooperation and motivating participation.

ensures that these added players should differ only on religious identity and do not systematically differ on skill or motivation. A balance table comparing the 28 Muslim and 28 Christian players added to the teams shows that both groups are highly comparable across a variety of demographic dimensions (Figure 2). The research design thus contains three sources of randomness: (1) which teams are selected to join the tournament, (2) which of these teams are assigned to the treatment condition, and (3) which additional players are assigned to each team conditional on treatment status. Each team has a total of 14 players and one (Christian) coach, yielding a sample of $n = 210$. Half of these ($n = 105$) are treated. To integrate the new players, encourage buy-in, and build team identity, each team received new uniforms and attended a one-day team orientation.[7] Guidelines on substitutions in addition to a high-intensity format that requires regular rotations between players ensured roughly equal playing time between all team members.[8]

Most contact interventions rely on survey responses or lab-in-the-field behaviors measured on the same day of the intervention that can struggle with external validity (Paluck, Green and Green 2017). I instead rely on a combination of attitudinal and behavioral outcomes measured between one day and four months after the intervention to detect the persistence of any treatment effects. To allay concerns that the sample size is too small to trust asymptotic results, I conduct block-bootstrapped analyses to generate predicted probabilities (distributions presented in Figures 6 to 11) and standard errors clustered at the team level ($n = 14$ clusters), in addition to showing that the

---

[7]Orientations for treated and control teams were held separately in order to avoid contact spillovers between the two conditions.

[8]The teams play 6-a-side, 60-minute matches on a half-pitch, the common format among amateur teams in this area. The intensity of this set-up means that players must alternate regularly. Coaches are instructed to ensure that all team members should play for roughly the same number of minutes each game. An additional constraint is that each player must sit out two games over the course of the league such that the substitute to starting player ratio is 1:1 in order to accommodate the 6-a-side format. Muslim players therefore played roughly the same number of minutes as their Christian teammates.

results hold with a permutation test (Table 2). I use two strategies to increase efficiency and precision for attitudinal outcomes. First, all participants take a baseline (T1) survey the day before the league begins in addition to a follow-up survey (T2) on the last day of the league, with a 100% response rate for both. The survey covers demographics, well-being, views on Iraqi society, beliefs regarding Iraq's future, and displacement experience at the hands of ISIS asked last so as to avoid negative priming or ordering effects (McFarland 1981). I take several T1 items as covariates in the main analyses to increase precision given the small sample size.[9]

Second, following Broockman, Kalla and Sekhon (2017), I combine similar survey items into an index in order to reduce measurement error. I do this by running an unsupervised hierarchical clustering algorithm on the T1 data collected for Christian respondents (n = 198) with a dummy treatment vector. Using a data-driven method to identify latent clusters in the survey data removes subjectivity from the process of manually choosing which items should form one index. The twenty-two survey items of interest were then collapsed into six indices, three of which serve as primary attitudinal outcomes. The "elbow" in the scree plot produced by this method (Figure 12 in the Appendix) plateaus sharply at three clusters, indicating that three primary indices should suffice. As a final step, I conduct a factor analysis on the six clusters to create scores that will serve as T2 outcomes.[10] The resulting indices – now dependent variables – cover trust toward Muslims, belief in coexistence, and the salience of Iraqi identity (Table 1), and align closely with theoretic expectations.[11] These indices are "primary" because they are most relevant to prejudice toward Muslims, while the three secondary indices cover freedoms, optimism, and generalized trust that are less likely to be affected by Muslim contact (Table 10). The items in these indices, along with T1 covariates, the estimation of average treatment effects (ATE) and heterogenous treatment effects, were pre-registered with EGAP before T2 outcomes were collected.[12]

I also measure three behavioral outcomes co-designed with local research staff, themselves displaced Iraqi Christians, and drawn from focus groups to capture organic and locally relevant manifestations of Muslim-Christian trust.[13] First, I test whether contact can bolster trust in non-sports settings. To do this, I invite all players to attend a dinner at sunset time on the 8th day of Ramadan (one week after the league). It is common knowledge that this

---

[9]These covariates include standard demographic traits, pre-displacement contact with Sunni Arabs, and abuse suffered at the hands of ISIS.

[10]If fewer than 50% of the variables in a given index for a given observation have missing values, I impute the missing values using medians. If over 50% of the variables in an index have missing values, I drop this observation.

[11]Items 2.1 and 2.2 are taken from the 2004 survey of Iraqis by Moaddel, Tessler and Inglehart (2008).

[12]EGAP registration #20170603AA.

[13]The same Iraqi research assistants record all outcomes in order to avoid pro-generous behavior encouraged by the presence of foreign observers (Cilliers, Dube and Siddiqi 2015).

Table 1: Primary Behavioral and Attitudinal Outcomes

Behavioral Outcomes

1. Attend Ramadan dinner event (three weeks post-intervention)

2. Train with Muslims at least once a week (four months post-intervention)

3. Register for mixed league in the future (one day post-intervention)

Attitudinal Indices

1. Trust toward Muslims

    1.1 Agree with adage that "even if a Muslim is a piece of gold, it would burn a hole in your pocket"
    1.2 Willing to sell land to a Muslim
    1.3 Generally speaking, agree that most people can be trusted
    1.4 Do not believe that Sunni Arab civilians are responsible for their suffering
    1.5 Believe that Sunni Arabs are welcoming toward Christians
    1.6 Believe that most Sunni Arabs did not approve of the actions of ISIS

2. Belief in Coexistence

    2.1 Believe that it is arbitrary to divide Iraqis in to ethnic and sectarian identities
    2.2 Believe that life would be better if Iraqis treated each other as Iraqis first
    2.3 Believe that tolerance is an important quality to teach children
    2.4 Believe that Christians need to arm themselves for protection
    2.5 Believe that life these days is unpredictable and dangerous

3. Salience of Iraqi Identity

    3.1 Identify above all else as Iraqi or Arab
    3.2 Proud or very proud to be Iraqi
    3.3 Believe that Iraqis of different backgrounds could live together peacefully

---

timing coincides with the fast-breaking meal (*iftar*) for Muslims. Players are encouraged to bring their families and friends, meaning that Christians are confronted with the option of socializing not only with Muslim players but with the latter's family members and friends as well. The outcome of interest is whether a player attends this mixed social event, and conditional on attendance, whether he brings his family – a more committed manifestation of social trust. For most if not all, this dinner was the first instance of inter-mingling since displacement three years prior.[14] Second, the T2 survey asks players if they would like to register for a mixed league in the future. Third, we telephone players four months after the league concludes to find out whether they regularly train with

---

[14]The lack of intergroup contact after displacement is due in part to self-segregation, but is largely explained by security checkpoints outside Muslim and Christian camp areas prohibiting access to outgroups.

Muslims. Having introduced participants to a practice field in neutral territory accessible to members of both groups, it is feasible that Christians in the control and treatment conditions could opt in to training with Muslims (encountered within the league or recruited from the broader community) after the intervention. Analyses are limited to Christians players ($n = 168$) as all Muslims are, by definition, in the treatment group.[15]

Results

Beginning with the three primary attitudinal outcomes, I find that the coexistence index yields a statistically significant treatment effect of 0.31 standard deviations (Table 2). This index consists of five items relating to the arbitrariness of sectarianism, the importance of tolerance, and perceived safety of life "these days" including the need for Christians to arm themselves for protection (Table 1). This effect suggests that Christians on mixed teams were more likely not only to denounce sectarian divisions, but also to perceive their environment as safer relative to those on homogenous squads. In fact, when parsing out the item on perceived security, Christians were around 33% more likely to report feeling safe relative to those in the control condition (p-value = 0.031, Table 15).[16] Statistically significant treatment effects when breaking down the indices into their component parts further highlights their validity from a measurement perspective. No statistically significant treatment effects were observed for the other two indices measuring trust toward Muslims and the salience of Iraqi identity, however.

The effects of contact extend into lasting behaviors. Table 2 illustrates predicted probabilities adjusting for covariates. Having Muslim teammates increases the probability of attending the Ramadan dinner by 26.1 percentage points (77.7% vs. 51.6%).[17] Intergroup contact therefore increased voluntary socialization by around 50% at a p-value of 0.08 (row 1 of Table 2). Bringing ones' family to an event open to Muslims is a particularly high bar for trust, which helps explain why only six players were accompanied by their families. Of these six players, four were from the treated group. Despite the progress that had been made in closing the trust gap between the two groups, most players were not ready to bring their families to a mixed event and a few shared this concern with

---

[15]Coaches were excluded from the analysis first because contact between coach and player is fundamentally different from contact between teammates, and second, because coaches have a vested interest in attending all league-related events (our behavioral outcomes) in order to demonstrate reliability and secure their teams' place in future competitions. I do however compare T1 and T2 survey responses among Muslim players ($n = 28$) for an indication of shifts in Muslim tolerance toward Christians (Table 13).

[16]Linear ATE regression with the same covariates reported in Table 8 with a binary outcome, agreeing or disagreeing with the statement that "life these days is unpredictable and dangerous."

[17]The raw means align with these predictions at 52% and 70% attendance, respectively (Table 7). The slight difference between predicted and raw means can be explained by the importance of adjusting for covariates in small samples. The randomization process did produce balance across almost all covariates, however, as shown in Table 6.

Table 2: Main Results (*n* = 168)

	Control	Treated	p-value
Behavioral Outcomes			
Iftar Attendance	51.6% (29.8, 73.0)	77.7% (53.6, 98.5)	0.081
Mixed League Sign-Up	92.6% (87.7, 97.1)	97.4% (93.5, 99.8)	0.058
Train w/ Muslims	22.1% (-0.04, 49.4)	65.6% (32.1, 97.7)	0.089
Attitudinal Outcomes			
Belief in Coexistence	-0.03 (-0.32, 0.25)	0.31 (0.06, 0.59)	0.018
Iraqi Identity	-0.01 (-0.46, 0.44)	0.15 (-019, 0.50)	0.428
Muslim Trust	-0.07 (-0.37, 0.19)	0.03 (-0.17, 0.28)	0.432

Entries are block bootstrapped predicted probabilities drawn from 2,000 samples and based on an OLS model with demographic controls for age, number of children, whether one lives in a camp, education, home ownership, pre-displacement income, church attendance, abuse experienced by ISIS, and whether the player was an added or core team member. T1 Muslim prejudice controls are: agreement with the derogatory adage that "even if a Muslim is a piece of gold, [he] will burn a hole in your pocket", how welcoming one finds Sunni Arabs, and how one characterizes the relationship between Christians and Muslims before the 2003 war. Only this final prejudice variable is included for for the attitudinal outcomes, as items in these indices are co-linear with T1 trust items. All variables coded in a pro-trust direction. Missing data imputed using medians. Block bootstrapped standard errors are clustered at the team level.

the research staff. Participants on mixed teams were also more likely to sign up for a mixed league in the future although both groups signed up at high rates (92.6% vs. 97.4%, Table 2). The high baselines here are attributable to the wording of the question: "Would you like to sign up for another league like this one in the future?" Almost all participants (92%) reported a "very good" or "excellent" experience with the league, implying that this outcome captures satisfaction with the overall tournament rather than trust toward Muslims. Uniformly positive evaluations of the league also render it difficult to test *H4*, that positive contact is better suited for trust-building than negative contact. Lastly, the predicted probability of training with Muslim players (encountered in the league and recruited from the broader Muslim community) four months after the intervention is 22.1% for control units compared with an over three-fold jump to 65.6% for treated units (Table 2).

Discussion: Incentives, Cooperation, and the Micro-Foundations of Social Trust

As predicted by Putnam, the league allowed Muslims to build, refine, and transmit good reputations that soon became apparent to Christians across all teams (Putnam, Leonardi and Nanetti 1994). The trusted flow of information between Christian participants diffused the knowledge that the Muslim players were trustworthy, cooperative, and unthreatening. Observations from local research staff attest to intense prejudice at the league's beginning gradually giving way to respect. In the league's early days, Christian players informed staff that "you can't trust Muslims," "we don't want Muslims, they will ruin the league," and that "we don't want them coming to our field" despite the fact the field was a public space located in the Ankawa city center. One coach walked out of the recruitment meeting and threatened to pull his team if Muslims were included. Research assistants noted a tense orientation session where Christians insisted on speaking Neo-Assyrian Aramaic (unintelligible to the Arabic-speaking Muslim players) and did not introduce themselves. When a Muslim player arrived a few minutes late to a warm-up session, Christian players grumbled that "this is why you can't rely on [Muslims]" despite their group's own casual attitude toward punctuality. Feeling unwelcome, Muslim players sat as far as possible from their teammates when seated on the sidelines. The league proceeded in this divided manner for around three weeks. Paluck and Green's meta-analysis of the contact theory affirms that "naturalistic studies" are likely to involve some amount of negative contact experiences, such as misunderstandings or outright conflict that could negatively affect outcomes (Paluck, Green and Green 2017). That salient incentives – new uniforms, a professionalized league, and a trophy – could overcome the first order problem of antagonistic groups unwilling to interact with one another suggests that the Christians in this case were "potential integrators."

Camaraderie slowly began to build and norms condemning prejudiced speech and behavior emerged. The coaches began speaking Arabic during half-time conversations and the players followed suit, affirming Allport's recommendation that intergroup contact be condoned by communal authority figures (1979). Muslim players sat closer to their teammates when on the bench and pro-tolerant shifts were recorded in their T2 survey responses.[18] Four Muslim players interviewed after the league shared that they felt "uncomfortable at first" as newcomers but eventually went on to forge "many friendships." One interviewee even stated that he preferred to play with Christians

---

[18]I look at survey items making up the trust indices that were posed to both Muslims and Christians. Of the seven eligible items, five demonstrated shifts in a pro-tolerant direction when comparing T1 and T2 surveys in difference-in-means t-tests (Table 13). A sample size of 28 renders these differences statistically insignificant though indicative.

as they seemed less aggressive than his colleagues in the Muslim league.[19] By the fifth week, when Muslims on the *Annunciation* team confessed that taxi rides from their camp to the field were getting unmanageably expensive, their Christian teammates shared the $35 cost twice a week without hesitation despite a mean household income between $500 – $1,000 per month.[20] Group-level norms condoning derogatory speech about Muslims appeared to have shifted as a result of contact – highlighting normative changes as another path wherein civic organizations can create social capital (Christ et al. 2014). When one league coordinator joked in the penultimate week that it "wasn't so bad having Muslims after all," a Christian player replied, "why do you have to think in such sectarian terms all the time, come on, don't talk like that!" The team captains appeared similarly changed, proposing that all-Muslim teams could be brought into future rounds of the league and that all teams should be integrated. Another sign of prejudice reduction came in the league's final days, which coincided with the UEFA Champion's League final on June 3, 2017. Christian players on treated teams invited their Muslim teammates to watch the game at a local cafe with no prompting from the research staff. One research assistant commented at the prize-giving ceremony: "without this [league], none of these guys would have even looked at a Muslim... I think we re-wrote a small piece of this history."[21] During informal interviews with 20 Christian participants seven months after the league ended, almost all expressed optimism about moving back to their hometowns and refrained from the generalized anti-Muslim speech that previously dominated conversations with research staff.[22]

Control teams were also exposed to Muslims as fellow league participants and the spread of trustworthy reputations. Results from the final behavioral analysis show that two of the seven control teams were training with Muslim players four months after the intervention (Table 3), which involved recruiting new Muslim players from the local community in addition to practising with Muslim players encountered in the league. That any control teams took the initiative to contact Muslim players outside the experiment and arrange regular practice sessions is indicative of increased tolerance as a result of secondary contact. This outcome is mechanistically biased in favor of treated teams that absorbed Muslim players during the league, but remains informative given that treated teams were established, competitive squads in their own right that did not need to retain added players. As one

---

[19]The intergroup contact literature suggests that intergroup contact affects members of majority/high status groups and members of minority/low status groups differently (Pettigrew and Tropp 2006). While Muslims in Iraq are the majority/high status group in Iraq, the Muslim study participants are displaced and living in a predominantly Christian area, skewing conventional categorizations of status. This research is hence agnostic with regards to differential effects for low vs. high status groups, which would also require a larger sample of Muslims to detect group-level effects.

[20]This was the same team that expressed regret at not being able to continue training together four months after the intervention due to transportation barriers (Table 3).

[21]Interview, Kellsey Beal, July 3, 2017. Ankawa, Iraq.

[22]Christian participants in T1 even expressed resentment toward Muslim groups in no way affiliated with ISIS, such as Iraqi Shi'ites, who themselves were targeted by ISIS.

would expect, contact effects are intensified among those with Muslim teammates: five treated teams opted to train with Muslims, with an additional team expressing their desire to do so but citing the distance between the Muslim and Christian camps as prohibitive to regular practice. Although not statistically significant, most of the attitudinal outcomes (four-fifths of the individual items and two-thirds of the indices) show a pro-tolerant shift from T1 to T2 for control participants (Figure 5), although control participants were 25 percent of a standard deviation more likely to register negative attitudes on the coexistence index in T2 relative to T1 (Table 7).[23] Taken together, the mixed to mildly positive results among control participants support both the secondary (*H2a*) and adversarial (*H2b*) contact hypotheses: secondary contact can build trust on the whole, but cooperative contact is better suited to reducing prejudice than adversarial contact, in line with Lowe (2017)'s findings on the positive returns to collaborative contact during an inter-caste cricket league in India.

Can the weak but extant trust captured in some T2 outcomes among control units be attributed to secondary contact at the league level? A limitation of this study is that there is no pure control group on which to test the counterfactual of no contact with Muslims whatsoever. Representative surveys of displaced Christians in Iraqi Kurdistan collected within three weeks of the league's start and end are indicative, however. I find either negative or insignificant changes in prejudice toward Muslims among the general population over the duration of the league (Figure 4). Christians in this region are largely segregated from Muslims, tentatively suggesting that we should not expect to see spikes in tolerance among those with no Muslim contact. Even the wives and mothers of players on mixed teams, when interviewed, continued to express anti-Muslim opinions.[24] Given the extreme distrust toward Muslims even compared with other displaced Christians (Figure 3), residential segregation of Muslims and Christians in Ankawa, and the initial hostility toward including Muslims in the league, it seems plausible that players on all-Christian teams would not have signed up for a mixed league in the future (92.6%) or approached Muslims in the broader community to arrange weekly training sessions in neutral territory (22.1%) were it not for the contact facilitated by the tournament. If we take the general population as a crude "pure control" group and the control group as a placebo, changes in trust seem to vary with the extent of contact: no or negative attitudinal

---

[23]An OLS model regressing the coexistence index on individual fixed effects and time period yields a period coefficient of –0.25 (p-value = 0.04) on the sample of control participants (n = 98), with block-bootstrapped standard errors clustered at the team level.

[24]The four female relatives interviewed shared that they were comfortable with a mixed league because they trusted the institution running it, a local Christian relief group, in keeping with Allport's stipulation of approval by communal authorities as a necessary condition for contact to reduce prejudice (Allport 1979). The same interviewees also stated that they did not trust Muslim men overall but were pleased with the therapeutic nature of league participation for their husbands and sons. Christian women have no contact whatsoever with Muslim men in their daily lives, perhaps rendering the trust gap too large to overcome with secondary contact. Spillover effects within the family appear mild to non-existent in this case. Contact interventions aimed at women from different ethnic and religious groups thus warrant exploration, especially in conflict areas.

Table 3: Behavioral Changes at 4 Months

Team	Practice Status	Train with Muslims?
Control Teams		
The Dreams	Train weekly	✗
The Kings	Train weekly w/ the Annunciation	✓
Bakhdeda's Brothers	Train twice a week or more w/ Cubs & 2 Muslims	✓
Bakhdeda's Youth	Train weekly	✗
The Knights	Disbanded	✗
Peace Youth	Disbanded	✗
Nishtiman	Train weekly	✗
Treated Teams		
Ozal City Youth	"Friends" w/ 2 Muslims, train weekly	✓
Bakhdeda's Lions	Train twice a week or more, two Christians emigrated	✓
Guards of Nineveh Plains	Train weekly	✓
Happiness	Train weekly	✓
Bakhdeda's Lion Cubs	Train weekly w/ 2 Muslims and Bakhdeda's Brothers	✓
Bartella's Youth	Would start training for another league	✗
The Annunciation	Wanted to train but Muslims camps too far	✗

changes for those outside the league, modest shifts for those in the league, and the strongest effects among those with Muslim teammates in the league. Chalking up trusting behavior among the control group in T2 to within-league contact is observational, but speaks both to the collinear relationship between trust and contact, and the power of even secondary contact to foster social capital through light-touch exposure to the outgroup.

Overall, intergroup contact against the backdrop of conflict did not lead to misunderstandings or even violence in ways cautioned by the contact literature (Paluck, Green and Green 2017) or outlined in the adversarial contact hypothesis. All-Christian teams did not play more aggressively, for instance, when facing off against mixed teams with Muslim players in their ranks. I investigate these dynamics using match-level data on red cards, yellow cards, and referee identifiers. If there is systematic aggression between Christians on control teams and Muslims on treated teams, then we should see a higher number of infractions, on average, in fixtures that bring together a treated

Table 4: Opposition Team Identity and Match-Level Aggression

	Dependent Variable		
	Total Cards	Yellow	Red
	(1)	(2)	(3)
Intercept	2.45	2.51**	−0.06
	(1.52)	(1.25)	(0.46)
Mixed	0.19	0.15	0.04
	(0.56)	(0.56)	(0.23)
Both Treated	0.07	−0.08	0.15
	(0.69)	(0.69)	(0.28)
Referee F.E. & Goals	✓	✓	✓
Observations	98	98	98
R^2	0.13	0.15	0.06
Adjusted R^2	0.02	0.04	−0.06

Note:	*p<0.1; **p<0.05; ***p<0.01

team with a control team than matches between teams in the same treatment condition. Table 4 demonstrates that there is no statistically significant difference in the number of cards doled out in matches between two Christian (control) teams, two Muslim-Christian (treated) teams, and two mixed (treated-control) teams. Grouping together teams of the same treatment status (both control or both treated) and comparing them with mixed match-ups similarly fails to yield a statistically significant difference in means (Table 5). It thus appears that the gains to social trust did not come with downside of increased aggression among participants in the experimental eco-system. Secondary contact with Muslims on other teams, at the very least, did not lead to outright conflict among control participants, building on scholarship illustrating that secondary contact can increase empathy and perspective-taking in divided societies (Simonivitz, Kezdi and Kardos 2017; Lustig 2003; Green and Brock 2002).

How were participants able to avoid the potentially negative effects of contact? Theories of intergroup contact during conflict as well as observational studies of interethnic peacebuilding programs suggest that non-compliance

and even violence are commonplace (Pettigrew 2016; Paluck, Green and Green 2017). Observational and experimental evidence reveal that both Muslim and Christian players contained potential conflict by censuring aggressive speech and actions toward outgroup members – an ingroup policing strategy to prevent disagreements from escalating (Fearon and Laitin 1996; Ditlmann and Samii 2016; Cikara and Paluck 2013). I leverage the as-good-as random assignment of Muslim referees affiliated with the sports facility in order to test the effects of the referee's religious identity on player aggression. I use the number of yellow and red cards shown by the referee as a proxy for match-level aggression, as cards serve to caution a player for unsporting behavior or rules infractions.

The 6% of games officiated by a Muslim referee saw an average of 1.80 cards given out, as compared with an average of 3.03 cards for games overseen by Christian officials (p-value = 0.10). Anecdotally, the usually expressive Christian players did not dispute any of the Muslim referees' decisions. A few Christian players even commented that they wanted to hire a Muslim referee viewed as "very fair" for future fixtures. Likewise, Muslim players who did not typically dispute the calls of Christian referees felt free to voice their opposition only with the Muslim referees. The same approach was observed with fellow players: Christians and Muslims appeared comfortable with challenging players – sometimes aggressively – from their own religious group, but seemed to deliberately avoid confronting those from the "other" group. Social psychologists consider ingroup censuring as a cooperative strategy associated with improved outgroup regard (Avenanti, Sirigu and Aglioti 2010). It would be a rational decision for the out-numbered Muslim players to avoid conflict even with no ingroup censuring at play. The restraint excercized by Christian players, however, syncs well with an explanation privileging greater empathy toward outgroup members experiencing harm – in this case, Muslims being targeted verbally or physically in ways that could be construed as intolerant given the political context. This qualitative evidence suggests that Christian teams were not more likely to target their aggressive behavior at Muslim players.

These decreases in discriminatory behavior, norms, and speech seem to map more clearly onto the behavioral outcomes than the attitudinal indices, however. In accordance with experimental evidence on inter-religious contact in Nigeria, I similarly find that social contact is more likely to change prejudiced behavior rather than self-reported attitudes (Scacco and Warren 2016). One explanation is social desirability bias. Attaining a tolerant score on the trust index requires expressing favorable attitudes specifically toward Muslims in five out of six items, while a pro-trust score on the the Iraqi identity index necessitates foregoing attachment to a Christian identity that is has grown even more fierce in the wake of the existential threat to Iraq's Christian community during and after displacement (Table 1). The sample size may have been insufficiently powered to detect minor shifts in these

dimensions. The likely presence of social desirability bias further stresses the need to measure implicit behaviors in such environments. Taking these findings on their face, however, the null effects on identity but not belief in coexistence may shed light on the role of national identity in divided societies. At least in this case, contact did not shift group identities entrenched by conflict. Instead, contact boosted optimism that warring groups could live together in peace.

Heterogenous Effects

Are "moderate" Christians driving these effects or those most hostile toward Muslims to begin with? Although under-powered, an analysis of heterogenous treatment effects sheds light on this question while further illuminating the importance of intergroup contact, associational membership, and the interaction between the two. I record three measures of contact with Sunni Arabs before displacement: how many Sunni Arabs one counted among their friends, the rate at which respondents frequented Sunni Arab stores back home, and how many Sunni Arab neighbors one had before displacement. Treatment effects for *iftar* attendance roughly doubled (from around 14% to 30%) for those with few Sunni Arab friends and for those who rarely or never visited Sunni Arab businesses (Table 14). This pattern demonstrates that the treatment was most effective for those with the least prior contact with Sunni Arabs. Meaningful intergroup contact therefore provide first-hand experience and information regarding the "other" for those who come from insular and segregated settings. Facilitated by a constant stream of communication between Christian group members, information regarding the good reputations of Muslim players spread quickly.

To fully understand this picture, one should consider why some respondents have lower rates of previous contact with Muslims in the first place. A plausible explanation is that segregation correlates with prejudice or distrust. Indeed, those who held more prejudiced views of Muslims in the baseline survey experienced amplified treatment effects.[25] I find the opposite trend for those who report having Sunni neighbors before displacement – living next to Muslims is associated with smaller treatment effects (15% vs. 29%, Table 14). While this may seem counter-intuitive at first, it has become clear that ISIS was able to identify the houses of Christians by extracting information out of Sunnis in the neighborhood. Even if coerced, this collusion was seen as a betrayal by many Christians according to the most exhaustive report on the ethnic cleansing of the Nineveh Plains (Center for the

---

[25]Respondents who believed that all Sunni Arabs approve of ISIS are 5 percentage points more likely to attend the *iftar* than those who view these Muslims as innocent (28% vs. 22%, Table 14). Additionally, the third of Christian participants who did not believe that tolerance was an important value to teach children were likewise more likely to experience large treatment effects (27%, p-value = 0.09).

Prevention of Genocide 2016). The first Christian to return his home in the ISIS-emptied city of Bakheda told Public Radio International: "When I bump into [my Muslim neighbors] now, they turn their faces and walk away... They know what they did. They know they're guilty. I don't even say hello to them" (Hall 2017). Nearly four-fifths of Christian study participants are from Bakhdeda (78.6%). Christians with Muslim neighbors feel deeply betrayed, and treatment effects are dampened accordingly. These sub-group trends are consistent with the idea that intergroup contact is most effective for those with the least prior contact with the outgroup, and relatedly, those with prejudiced baseline attitudes. Those with the strongest feelings of distrust, however, were less affected – suggesting an upside-down U-shaped relationship between baseline distrust and the effectiveness of contact.[26]

Another heterogenous effect illustrates to the importance of civic organizations: those who report membership in other civic associations were 0.46 standard deviations (p-value = 0.10) more likely to affirm a belief in coexistence as measured by the relevant attitudinal index in an interactive model (Table 14). Just under a third of participants (29%) report being a member of a political party, professional association, or youth group – a slightly lower rate than the average for men in Iraq (33%) and the Arab world more broadly (41%) according to the latest 2010–2015 wave of the World Values survey (Association 2015).[27] Given the small sample size, clustered standard errors, and the power required to detect statistically significant interaction effects, this finding merits discussion. Of the respondents who report membership in another civic association, a third are members of a youth group, a fifth have joined a professional syndicate, and the remaining tenth are scattered between political parties, volunteer groups, and religious organizations. Almost all of these organizations tend to have religiously homogenous membership and are thus not cross-cutting. Others, particularly professional syndicates, qualify as cross-cutting as as a result of their inclusive purpose and membership. Importantly, none of these organizations are both religiously homoge-nous and exclusive in purpose, such as the Ku Klux Klan or Weimar-era clubs that ultimately undermine social capital (Tesmer et al. 1997; Berman 1997). The increase to treatment effects among those already other associa-tions indicates that civic culture is re-enforcing – the norms, skills, and habits encouraging broad cooperativeness in one (associational setting can be transferred and amplified in another (Putnam, Leonardi and Nanetti 1994). Keeping in mind that all of these teams were religiously homogenous groups before the intervention, the statis-

---

[26]These results were obtained by subsetting the data by level of contact (few to none vs. many) and reporting the statistically significant treatment effects. The null hypothesis here is that there is no effect among the subsetted group. Unfortunately, the sample size is too weak to formally compare both subsetted groups against each other, though the interaction term between the treatment dummy and a binary "frequency of visiting Sunni Arab shops back home" variable was weakly statistically significance at around p = 0.17 (Table 14).

[27]Arab world is here defined as the 12 Arab League states sampled in the 2010–2015 World Values Survey: Algeria, Egypt, Iraq, Jordan, Kuwait, Lebanon, Libya, Morocco, Palestine, Qatar, Tunisia, and Yemen.

Table 5: Mean Infractions by Match-Up Treatment Status ($n = 99$)

	Same Status	Different Status	p-value
Total Cards	3.00	2.92	0.89
Yellow Cards	2.68	2.56	0.80
Red Cards	0.32	0.36	0.82

tically and substantively large average treatment effect results likewise suggest that participation in homogenous groups can lay the groundwork for social trust upon the addition of intergroup contact.

Alternative Explanations and External Validity

In order to test the effects of inter-religious contact, the ideal experiment would randomly assign religion. In reality, Muslims and Christians differ on more than just religion. One factor that partly explains these effects is that the added Muslim players, pressured by the hostile environment they found themselves in, might have been more motivated to perform. Individual-level data shows that Muslims and Christians scored a similar number of goals per game (0.135 compared to 0.131, respectively) and that added players were balanced on self-reported health and fitness measures (Table 2). Yet treated teams fared better overall, scoring a total of 125 points to the treated teams' 137.[28] In contrast with research documenting a positive link between team performance and diversity in science and the private sector (Hunt, Layton and Prince 2015; Nielsen et al. 2017; Hong and Page 2004) but not in sports (Timmerman 2000), three of the top five most successful teams (including the ultimate winners) were mixed. Every Christian medalist from the treatment group (the first and third place winners) attended the *iftar* dinner whereas 20% to 90% of Christians on the remaining five treated teams skipped the event. Upon watching games where Muslim players excelled, Christians in the control condition remarked that they would have preferred to absorb the top-scoring Muslim players rather than fellow Christians. A functionalist explanation alone does not explain increased tolerance toward Muslims, however: the entire squad of the mixed *Happiness* team attended the dinner despite landing in the bottom half of the league, while a lone member of the treated *Bartella's Youth* squad was present despite this team finishing at a respectable fourth overall. Nonetheless, skill and success in team environments appears to foster positive affect and enhance intergroup contact.

---

[28]Teams gained three points for a win, one for a draw, and none for a loss. See Appendix: League Set up for more details.

To what extent are these results externally valid? A limitation of this study is that the sample is not randomly selected from the general population. Every stage of randomization – selecting the participating teams, selecting 56 players from other teams to be added, and assigning treatment status – is conditional on participants already being amateur soccer players. The results may have been different if individuals from Ankawa were randomly selected to join teams that constructed from scratch. Treatment effects in this scenario would presumably have been diluted, as such a league would include many non-athletes with no interest or stake in winning a soccer league. And yet leveraging existing civic organizations has several advantages. Designing interventions around organic, grassroots organizations ensures that contact is as naturalistic as possible while avoiding unrealistic contact scenarios and lab-in-the-field situations rarely encountered in everyday life (Paluck, Green and Green 2017). Putnam similarly warns against the ineffectiveness of "transplanted" civil society groups, where attrition would likely be problematic (Putnam, Leonardi and Nanetti 1994). Even within the segregated setting of Ankawa, intergroup contact within the context of soccer did not raise eyebrows: Muslim participants interviewed four months after the intervention had no inkling that their invitation to join was part of a diversity mandate. Lastly, drawing on existing civic organizations and working with members who self-select into these settings renders the results externally valid to this sub-population of displaced males and thus maximally relevant to policy.

An aim of this study is to demonstrate that scholars can take self-selection into civic associations for granted while thinking creatively about how to transform homogenous groups into weakly cross-cutting ones. The experimental findings on intergroup contact and observational findings on positive spillover effects among control participants should give us reason for optimism: injecting intergroup contact into associational settings by virtue of competing in the same team or tournament can reap the trust-building rewards associated with cross-cutting groups. More-over, building on rather than circumventing locally resonant associations may lead to spillover effects within the community. To this end, the league gained substantial local traction. Hundreds of spectators from mothers and children to Peshmerga officers lined the field at every game and a Facebook group announcing fixtures, reporting scores, and streaming matches live attracted over 1,000 members (Figure 1).[29] In addition, coaches interviewed at the *iftar* event suggested that future iterations of the league either include teams from the Muslim amateur league as competitors or integrate Muslim players onto more Christian teams. This project therefore serves as a proof-

---

[29]At the time of this writing in November 2017, the page has 1,123 followers, a 5/5 rating, over 700 match photos (many uploaded by fans), and some videos viewed over 3,000 times. Roughly 300 – 400 spectators lined the field for every game; mothers and children, Peshmerga officers, and curious residents.

of-concept that unimposing intergroup contact can be embedded into homogenous organizations with inclusive, non-identity based aims – with positive consequences for social trust.

The question remains, however, of the extent to which these results are generalizable. Male athletes may be a distinct population for many reasons. For instance, attachment to one's identity as an athlete has been found to correlate with increased ethno-centrism in some contexts (Gatz, Messner and Ball-Rokeach 2002) but enhanced cooperativeness in others (Duda 1989). To address this question, I compare baseline pre-survey responses (T1) with responses from a representative survey (n = 1,115) of displaced Christians in Erbil. Figure 3 shows that the study participants differ from the general population in ways we would expect: the soccer players are younger, wealthier, more likely to be single, and have fewer children than other displaced Christians in the area. Experimental subjects differ on attitudinal dimensions, too – they are consistently more prejudiced toward Muslims than their counterparts from the general population. The low levels of baseline trust toward Muslims among participants as evidenced by the T1 means in Figure 3 suggest that this sample of individuals is more hostile toward Muslims than the population from which they are drawn. In absolute terms, this hostility is reflected in the finding that 38.1% of participants reported that Christians should arm themselves for protection against Muslims in the T1 survey, with a similar portion (35.1%) agreeing with the prejudiced folk saying that "even if a Muslim is made of gold, [he] will burn a hole in your pocket." It is possible that Christians consider the subset of Muslims included in this study as less complicit in the ISIS raids relative to Muslims who chose to remain in ISIS-occupied territories. On the other hand, at least a third of Muslim respondents reported fleeing their homes after the ISIS invasion.[30] A similar selection concern is that the minority of Christians who remained in Iraq after 2003 might be committed to preserving a Christian presence in Iraq and thus particularly ethno-centric. The T1 survey data show that this is not the case, with 68.9% of Christians reporting that they would emigrate from Iraq given the chance. Lastly, sharing a (nominal) ethno-religious affiliation with ISIS also renders Sunni Arabs are a particularly unwelcome Muslim group compared with Kurds or Shi'ite Arabs, for instance.

The political context around the intervention's start time would also bias results toward null or negative treatment effects. During focus groups run in October 2016, Christians were optimistic that they could return to their home-towns in the Nineveh plains area following liberation. A week before the league launched, Christian participants became aware of the extent of the damage inflicted on their homes. No governmental or civil society actor was

---

[30]28.2% of Muslims in the sample fled their homes after the summer of 2014, which is potentially a lower bound estimate given that the same proportion did not specify the month of displacement.

willing or able to help with the reconstruction effort beyond supporting fewer than 100 families, leaving the victims – drained of savings two years after displacement – to rebuild their homes.[31] More distressing news came in early August before the final behavioral outcome of training with Muslim players was collected. The Assyrian church announced they it would be forced to cut off rent subsidies equalling around $400 per month to displaced Christian families, who have an average income of $500 – $1,000 per month. Around the same time, the Kurdish government declared that public schools accessible to displaced children would be relocated almost 80 miles away from Ankawa in the city of Qaraqosh. In short, the positive effects of intergroup contact reported here are particularly difficult to attain given the challenging political environment that displaced Christians inhabit.

Why did Christians participate in a mixed league given the distrust of Muslims and availability of outside options in the form of local Christian leagues? I overcame these constraints by offering locally relevant incentives to join the league, including new uniforms, reserving and paying for the field, and hiring professional referees. These incentives brought captains back to the table during the league's inception meeting, where they then proposed trophies to honor the top three teams and subsequently agreed to partake in the league. Professionalizing the league in a way that resonated with participants partly explains the low dropout rates and perfect compliance. If incentives were not needed to engage in intergroup activities, then participants are arguably tolerant enough such that trust interventions are not particularly necessary nor effective. In conflict settings where the stakes of trust interventions are high, incentives designed with local partners are important for encouraging participation. Such incentives are needed to overcome the psycho-social barriers to participating in intergroup activities facing potential integrators.

Conclusion

An ongoing refugee crisis, global tides of populism, and worsening sectarianism in the Middle East have reinvigorated the longstanding questions of whether intergroup contact can reduce prejudice and how social trust can be bolstered. Disparate literatures have highlighted the importance of mutual cooperation and a shared goal (Allport 1979; Fukuyama 2001; Diamond 1993), equal power statuses during contact (Allport 1979; Pettigrew 2016), and the role of norms encouraging reciprocity and admonishing opportunistic behavior (Putnam, Leonardi and Nanetti 1994; Fukuyama 2001) in building social trust. Sports settings are remarkable for their ability to

---

[31]The cost of reconstruction is estimated to be between $50 to $100 billion (Stasser 2017). Christians in the study reported having spent an average of 50 – 75% of their savings when surveyed in March 2017.

combine all of these conditions as well as their ubiquity across the globe.[32] Building on the illustrative example of sports clubs, I show that combining cross-cutting interests with meaningful intergroup contact and salient incentives can create the conditions necessary for prejudice reduction over time. A clear stake in cooperation encourages committed participation and allows the trust-building potential of social contact to unfold naturally, suggesting that investments in civic culture should be targeted at associations that explicitly cut across social cleavages and whose purpose is based on skill or interest rather than immutable identities (Boix and Posner 1996).

Civic cultures conducive to building and reinforcing the trust needed for political and economic development have been characterized as path dependent and sticky equilibria resistant to public policy changes (North 1990; Zamagni 1978; Putnam, Leonardi and Nanetti 1994; Sabetti 1996; Fukuyama 2001). This study demonstrates that cross-cutting groups can break the trap of "uncivic" communities. Intense baseline distrust, the trauma of displacement at the hands of the outgroup, and a charged political climate would make this case a hard one for creating trust through intergroup contact. Moreover, prejudice reduction is thought to be particularly difficult when issue salience is high (Broockman, Kalla and Sekhon 2017), as it is with the sectarian Iraqi conflict. Despite these constraints, I have shown that intergroup contact within grassroots associations that cut across social cleavages – in this case, religiously mixed soccer clubs – persistently improved trusting attitudes and behaviors between displaced Christians and Muslims in Iraq. The results support the idea that weak ties such as "shared membership in a secondary association" (Putnam, Leonardi and Nanetti 1994) can create and sustain cohesion in spite of the social trust deficit left in the wake of war (Granovetter 1973; Banfield 1967). Civil conflicts often give rise to ethnocentric, prejudiced, and fear-driven beliefs of the other, which displaced people can carry with them in ways that stoke new conflict (Salehyan and Gleditsch 2006). With the right incentives, outgroup beliefs – and by extension, uncivic cultures – are reversible.

Those who had barely interacted with Sunni Arabs in the past – and relatedly, those who most distrusted Muslims – saw the greatest gains in tolerant attitudes and behaviors on average. Qualitative evidence depicts how the trusted flow of communication within the league setting was able to facilitate Muslims building and transmitting good reputations, thus filling information gaps on the nature of Muslims for those with little prior contact. I also find observational evidence for the emergence of norms against prejudiced speech and the containment of potentially aggressive behavior through ingroup censuring. Taken together, this supplementary evidence suggests

---

[32]According to the 2010-2015 wave of the World Values Survey, about 22% of the world's population report membership in a sports association (Association 2015).

that meaningful intergroup contact can reduce prejudice by sharing reputational information, creative cooperative norms, and encouraging ingroup policing against the backdrop of tense political climates.

Even taking self-selection into associations for granted, encouraging diverse membership within existing associations in unimposing ways can create meaningful social contact. Modest increases in tolerant opinions recorded among control participants indicate the potential for secondary contact to foster trust, and at the very least show that homogenous organizations common to "uncivic" societies do not preclude social trust under the right conditions. As I have argued here, there is a case to be made for building interventions on top of existing associations and foregoing some benefits to generalizability in order to obtain the most policy-relevant results possible. Co-designing peacebuilding programs with indigenous NGOs has proven particularly fruitful in Iraq, where these organizations brokered peace accords that prevented revenge bloodshed after ISIS massacres and paved the way for the return of 380,000 displaced Iraqis – saving the U.S. $150 million per month (Lindborg 2017). More broadly, the findings on integrating civic organizations can inform the $612 million allotted by USAID for civil society, conflict mitigation, and reconciliation activities in 2016 (USAID 2017), millions of dollars spend on reconciliation programs focused on "exposure to the other" (McKone 2015), and FIFA's $4.3 million annual commitment to social development programming (FIFA 2016). Nevertheless, future work should manipulate the conditions of contact, test specific mechanisms of prejudice reduction, investigate whether new, constructed organizations are as conducive to building trust as membership in existing associations, and disentangle potential "friendship effects" toward intervention participants from generalizable out-group trust.

References

Al-Ameen, Shaida. 2014. "Christians displaced again by Iraqi violence.".

Allport, Gordon W. 1979. *The nature of prejudice*. Basic books.

Association, WVS. 2015. "World values survey." *Wave* 6:2010–2015.

Avenanti, Alessio, Angela Sirigu and Salvatore M Aglioti. 2010. "Racial bias reduces empathic sensorimotor resonance with other-race pain." *Current Biology* 20(11):1018–1022.

Banfield, Edward C. 1967. "The moral basis of a backward society.".

Bar-Tal, Daniel. 2000. "From intractable conflict through conflict resolution to reconciliation: Psychological analysis." *Political Psychology* 21(2):351–365.

Basu, Moni. 2016. "In biblical lands of Iraq, Christianity in peril after ISIS." *CNN* .

Bellin, Eva. 2012. "Reconsidering the robustness of authoritarianism in the Middle East: Lessons from the Arab Spring." *Comparative Politics* 44(2):127–149.

Bellows, John and Edward Miguel. 2009. "War and local collective action in Sierra Leone." *Journal of public Economics* 93(11):1144–1157.

Berman, Sheri. 1997. "Civil society and the collapse of the Weimar Republic." *World politics* 49(03):401–429.

Blattman, Christopher. 2009. "From violence to voting: War and political participation in Uganda." *American political Science review* 103(02):231–247.

Boix, Carles and Daniel N Posner. 1996. *Making Social Capital Work: A Review of Robert Putnam's Making Democracy Work, Civic Traditions in Modern Italy*. Number 96 Center for International Affairs, Harvard University Cambridge, MA.

Broockman, David E, Joshua L Kalla and Jasjeet S Sekhon. 2017. "The Design of Field Experiments With Survey Outcomes: A Framework for Selecting More Efficient, Robust, and Ethical Designs." *SSRN* .

Brubaker, Rogers and Frederick Cooper. 2000. "Beyond ?identity?" *Theory and society* 29(1):1–47.

Burns, Justine, Lucia Corno and Eliana La Ferrara. 2015. Interaction, prejudice and performance. Evidence from South Africa. Technical report Working paper.

Cameron, Lindsey, Adam Rutland, Rupert Brown and Rebecca Douch. 2006. "Changing children's intergroup attitudes toward refugees: Testing different models of extended contact." *Child development* 77(5):1208–1219.

Center for the Prevention of Genocide, Simon Skjodt. 2016. "'Our Generation is Gone': The Islamic State's Targeting of Iraqi Minorities in Ninewa [Bearing Witness Report].".

Christ, Oliver, Katharina Schmid, Simon Lolliot, Hermann Swart, Dietlind Stolle, Nicole Tausch, Ananthi Al Ramiah, Ulrich Wagner, Steven Vertovec and Miles Hewstone. 2014. "Contextual effect of positive intergroup contact on outgroup prejudice." *Proceedings of the National Academy of Sciences* 111(11):3996–4000.

Cikara, Mina and Elizabeth Levy Paluck. 2013. "When going along gets you nowhere and the upside of conflict behaviors." *Social and Personality Psychology Compass* 7(8):559–571.

Cilliers, Jacobus, Oeindrila Dube and Bilal Siddiqi. 2015. "The white-man effect: How foreigner presence affects behavior in experiments." *Journal of Economic Behavior & Organization* 118:397–414.

Collier, Jane and Rafael Esteban. 2007. "Corporate social responsibility and employee commitment." *Business ethics: A European review* 16(1):19–33.

Crisp, Richard J and Rhiannon N Turner. 2009. "Can imagined interactions produce positive perceptions?: Reducing prejudice through simulated social contact." *American Psychologist* 64(4):231.

Deutsch, Morton. 1977. *The resolution of conflict: Constructive and destructive processes*. Yale University Press.

Diamond, Larry. 1993. "Causes and effects." *Political culture and democracy in developing countries* pp. 411–435.

Ditlmann, Ruth K and Cyrus Samii. 2016. "Can intergroup contact affect ingroup dynamics? Insights from a field study with Jewish and Arab-Palestinian youth in Israel." *Peace and Conflict: Journal of Peace Psychology* 22(4):380.

Duda, Joan L. 1989. "Relationship between task and ego orientation and the perceived purpose of sport among high school athletes." *Journal of sport and exercise psychology* 11(3):318–335.

Enos, Ryan D. 2014. "Causal effect of intergroup contact on exclusionary attitudes." *Proceedings of the National Academy of Sciences* 111(10):3699–3704.

Fearon, James D and David D Laitin. 1996. "Explaining interethnic cooperation." *American political science review* 90(4):715–735.

Fehr, Ernst, Michael Kosfeld and Urs Fischbacher. 2005. "Neuroeconomic foundations of trust and social preferences.".

FIFA. 2016. "FIFA extends the reach of its global social development initiative.".

Forbes, Hugh Donald. 1997. *Ethnic conflict: Commerce, culture, and the contact hypothesis*. Yale University Press.

Fukuyama, Francis. 1995. *Trust: The social virtues and the creation of prosperity*. Number D10 301 c. 1/c. 2 Free Press Paperbacks.

Fukuyama, Francis. 2001. "Social capital, civil society and development." *Third world quarterly* 22(1):7–20.

Gatz, Margaret, Michael A Messner and Sandra Ball-Rokeach. 2002. *Paradoxes of youth and sport*. SUNY Press.

Graf, Sylvie, Stefania Paolini and Mark Rubin. 2014. "Negative intergroup contact is more influential, but positive intergroup contact is more common: Assessing contact prominence and contact prevalence in five Central European countries." *European Journal of Social Psychology* 44(6):536–547.

Granovetter, Mark S. 1973. "The strength of weak ties." *American journal of sociology* 78(6):1360–1380.

Green, Donald P and Rachel L Seher. 2003. "What role does prejudice play in ethnic conflict?" *Annual Review of Political Science* 6(1):509–531.

Green, Melanie C and Timothy C Brock. 2002. "In the mind's eye: Transportation-imagery model of narrative persuasion.".

Hall, Richard. 2017. "Iraqi Christians are slowly returning to their homes, wary of their neighbors.".

Hammack, Phillip L. 2006. "Identity, conflict, and coexistence: Life stories of Israeli and Palestinian adolescents." *Journal of Adolescent Research* 21(4):323–369.

Hewstone, Miles, Ed Cairns, Alberto Voci, Juergen Hamberger and Ulrike Niens. 2006. "Intergroup contact, forgiveness, and experience of ?The Troubles? in Northern Ireland." *Journal of Social Issues* 62(1):99–120.

Hong, Lu and Scott E Page. 2004. "Groups of diverse problem solvers can outperform groups of high-ability problem solvers." *Proceedings of the National Academy of Sciences of the United States of America* 101(46):16385–16389.

Hunt, Vivian, Dennis Layton and Sara Prince. 2015. "Diversity matters." *McKinsey & Company* 1.

Jamal, Amaney A. 2009. *Barriers to democracy: The other side of social capital in Palestine and the Arab world.* Princeton University Press.

Jha, Saumitra. 2013. "Trade, institutions, and ethnic tolerance: Evidence from South Asia." *American political Science review* 107(4):806–832.

Jha, Saumitra and Moses Shayo. 2017. "Valuing Peace: The Effects of Financial Market Exposure on Votes and Political Attitudes.".

Karlan, Dean S. 2007. "Social connections and group banking." *The Economic Journal* 117(517).

Knack, Stephen and Philip Keefer. 1997. "Does social capital have an economic payoff? A cross-country investigation." *The Quarterly journal of economics* 112(4):1251–1288.

Leitner, Michael J, Yair Galily and Pini Shimon. 2012. "The effects of Peres Center for Peace sports programs on the attitudes of Arab and Jewish Israeli Youth." *Leadership and Policy Quarterly* 1(2):109–121.

Lindborg, Nancy. 2017. "Stabilize Iraq After ISIS, Try a Method That Worked." *United States Institute of Peace* .

Litvak-Hirsch, Tal, Yair Galily and Michael Leitner. 2016. "Evaluating conflict mitigation and health improvement through soccer: a two-year study of Mifalot's "United Soccer for Peace" programme." *Soccer & Society* 17(2):209–224.

Lowe, Matt. 2017. "Types of Contact: A Field Experiment on Collaborative and Adversarial Caste Integration." *MIT Working Paper* .

Lugo, Luis, Alan Cooperman, James Bell, Erin O'Connell and Sandra Stencel. 2013. "The world's Muslims: Religion, politics and society." *World* .

Lustig, I. 2003. "The influence of studying foreign conflicts on students? perceptions of the Israeli Palestinian conflict." *Unpublished masters thesis, University of Haifa* .

McFarland, Sam G. 1981. "Effects of question order on survey responses." *Public Opinion Quarterly* 45(2):208–215.

McKone, Kelly. 2015. *Reconciliation in practice*. United States Institute of Peace.

Moaddel, Mansoor, Mark Tessler and Ronald Inglehart. 2008. "Foreign occupation and national pride: The case of Iraq." *Public Opinion Quarterly* 72(4):677–705.

Nielsen, Mathias Wullum, Sharla Alegria, Love Börjeson, Henry Etzkowitz, Holly J Falk-Krzesinski, Aparna Joshi, Erin Leahey, Laurel Smith-Doerr, Anita Williams Woolley and Londa Schiebinger. 2017. "Opinion: Gender diversity leads to better science." *Proceedings of the National Academy of Sciences* 114(8):1740–1742.

North, Douglass C. 1990. *Institutions, institutional change and economic performance*. Cambridge University Press.

Paluck, Elizabeth Levy, Seth Green and Donald P Green. 2017. "The Contact Hypothesis Revisited." *SSRN* .

Pettigrew, Thomas F. 2016. "In pursuit of three theories: Authoritarianism, relative deprivation, and intergroup contact." *Annual review of psychology* 67:1–21.

Pettigrew, Thomas F and Linda R Tropp. 2006. "A meta-analytic test of intergroup contact theory." *Journal of personality and social psychology* 90(5):751.

Posen, Barry R. 1993. "The security dilemma and ethnic conflict." *Survival* 35(1):27–47.

Putnam, Robert 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.

Putnam, Robert D, Robert Leonardi and Raffaella Y Nanetti. 1994. *Making democracy work: Civic traditions in modern Italy*. Princeton university press.

Rao, Gautam et al. 2013. "Familiarity does not breed contempt: Diversity, discrimination and generosity in Delhi schools." *Job Market Paper* .

Rohner, Dominic, Mathias Thoenig and Fabrizio Zilibotti. 2013. "Seeds of distrust: Conflict in Uganda." *Journal of Economic Growth* 18(3):217–252.

Rothstein, Bo and Eric M Uslaner. 2005. "All for all: Equality, corruption, and social trust." *World politics* 58(01):41–72.

Sabetti, Filippo. 1996. "Path dependency and civic culture: some lessons from Italy about interpreting social experiments." *Politics & Society* 24(1):19–44.

Salehyan, Idean and Kristian Skrede Gleditsch. 2006. "Refugees and the spread of civil war." *International Organization* 60(2):335–366.

Sands, Melissa L. 2017. "Exposure to inequality affects support for redistribution." *Proceedings of the National Academy of Sciences* p. 201615010.

Scacco, Alexandra and Shana S Warren. 2016. "Can Social Contact Reduce Prejudice and Discrimination? Evidence from a Field Experiment in Nigeria." *Unpublished working paper* .

Schroeder, Juliana and Jane L Risen. 2016. "Befriending the enemy: Outgroup friendship longitudinally predicts intergroup attitudes in a coexistence program for Israelis and Palestinians." *Group Processes & Intergroup Relations* 19(1):72–93.

Simonivitz, Gabor, Gabor Kezdi and Peter Kardos. 2017. "Seeing the World Through the Other's Eye: An Online Intervention Reducing Ethnic Prejudice." *American Political Science Review* pp. 1–8.

Stasser, Fred. 2017. "'U.S.-Light"in Iraq May Open Way for Russia, Iran." *United States Institute of Peace* .

Tajfel, Henri and John C Turner. 1979. "An integrative theory of intergroup conflict." *The social psychology of intergroup relations* 33(47):74.

Tesmer, John JG, David M Berman, Alfred G Gilman and Stephen R Sprang. 1997. "Structure of RGS4 bound to AlF 4–activated G iα1: stabilization of the transition state for GTP hydrolysis." *Cell* 89(2):251–261.

Timmerman, Thomas A. 2000. "Racial diversity, age diversity, interdependence, and team performance." *Small Group Research* 31(5):592–606.

UNHCR. 2017. "Iraq emergency." *UNHCR* .

USAID. 2017. "Shared Progress, Shared Future: Agency Financial Report, Fiscal Year 2016." *United States Agency for International Development* .

Varshney, Ashutosh. 2003. *Ethnic conflict and civic life: Hindus and Muslims in India.* Yale University Press.

Walter, Barbara F. 2004. "Does conflict beget conflict? Explaining recurring civil war." *Journal of Peace Research* 41(3):371–388.

Wood, Geof. 2003. "Staying secure, staying poor: the ?Faustian Bargain?" *World Development* 31(3):455–471.

Zamagni, Vera Negri. 1978. *Industrializzazione e squilibri regionali in Italia: bilancio dell'età giolittiana.* Mulino.

Appendix

League Set-Up

The teams play 6-a-side, 60-minute matches on a half-pitch, the common format among the amateur teams. The intensity of this set-up means that players must alternate regularly. Coaches are also instructed to ensure that players rotate regularly. Taken together, all team members should play for roughly the same number of minutes each game. An additional constraint is that each player must sit out two games over the course of the league such that the substitute to starting player ratio is 1:1, in order to accommodate the 6-a-side format.

Phase	Teams Included (N)	Start Date	End Date
One (round-robin	All (14)	March 27, 2017	May 20, 2017
Two (quarter-finals)	2nd—9th place teams (8)	May 23, 2017	May 30, 2017
Final Match	League vs. Play-off Champions (2)	June 2, 2017	June 2, 2017

The 99 matches played proceeded in two broad phases. Phase One is a classic round-robin set-up. Each team plays 13 matches over 8 weeks to reach a total of about 26 hours on the pitch per player when warm-up and cool-down time are taken into account. The team with the most points (three points for a win, one for a draw, and zero for a loss) secures a spot in the final while the bottom four are eliminated. Phase Two entails knock-out quarter-finals played by the remaining 2nd through 9th place teams. The winning team from Phase Two plays the winner from Phase 1 in the Final. The number of additional hours played by teams that qualify for the quarter finals ranges from two to ten.

Randomization and Balance

Figure 2 shows balance on social attitudes, self-reported health, and other background traits between the 28 added Christian and 28 added Muslim players. Balance on these indicators suggest that the religion of the added players does not correlate with observables that may affect their disposition or skill in ways that shape treatment effects. The hostile environment that Muslim players found themselves in may have motivated them to perform at a higher level than added Christian players, however, described in the Discussion section.

**Figure 2:** Balance Between Added Christian vs. Muslim Players ($n = 56$)
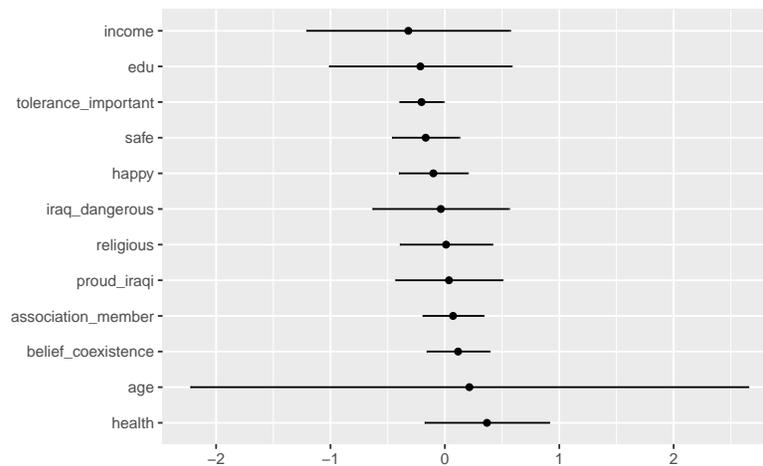
The two figures below speak to the generalizability of the findings. Figure 3 below compares the study sample with the general population from which they are drawn using a representative survey of displaced Christians in the Erbil area. Despite being around 11% more likely to believe that it is important to teach tolerance to children, Christian players have lower baseline levels of happiness, perceptions of safety, and trust toward Muslims, arguably making it harder to find positive treatment effects among this particular sub-group.

Figure 3: Comparison of Christian Participants ($n = 168$) and Displaced Christian Population ($n = 1,115$)
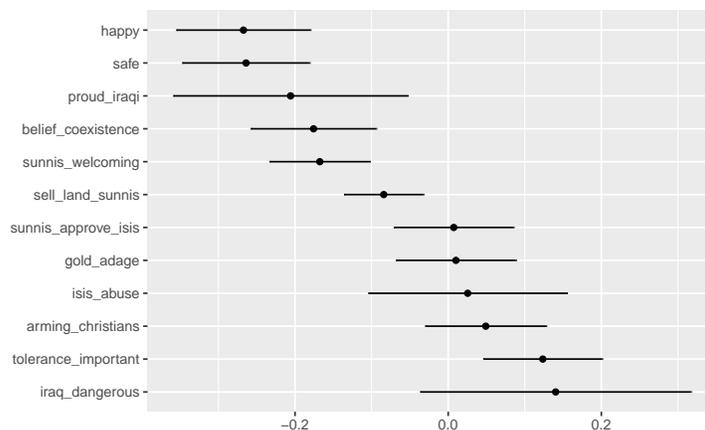
I make the case that although there is no pure control group in the study, the finding that a fifth of control participants were training with Muslims four months after the intervention can be attributed to secondary contact with Muslims within the league. The segregation of Erbil and low baseline attitudes toward Muslims suggests that Christians players would have been very unlikely to have meet Muslim players outside of the experiment. To support this claim, Figure 4 shows that there were no statistically significant changes in attitudes toward Muslims among the general population of displaced Christians in Erbil surveyed within three weeks of the league's start and end. Figure 5 points to a positive shift in tolerant attitudes when comparing point estimates for control units before and after the intervention, further suggestive evidence of secondary contact effects.

Figure 4: Prejudice among Displaced Christian Population in April/May vs. June ($n = 389$)
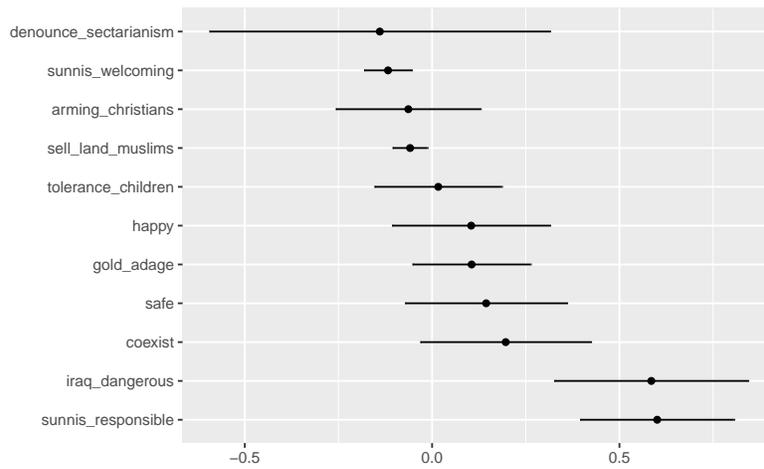
Figure 5: Shifts in Tolerant Attitudes among Control Group (T2 vs. T1) ($n = 98$)
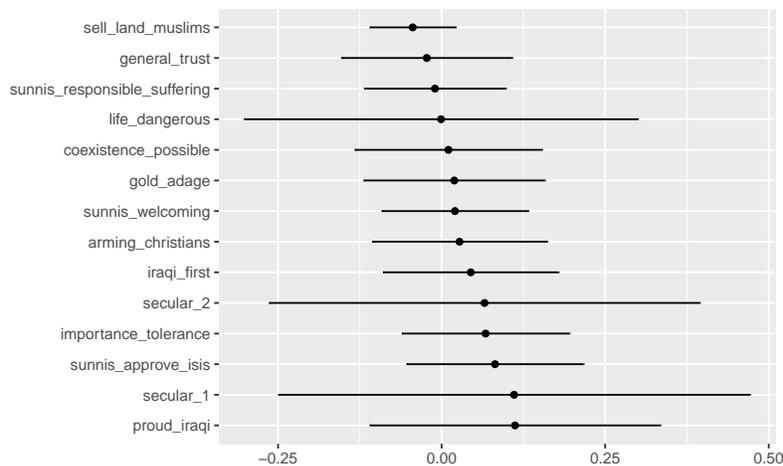
Table 6: Mean Balance Across Treated and Control Christian Players

	Control ($n = 98$)	Treated ($n = 70$)	Std. Mean Diff.	p-value
Demographics				
Age	22.0	21.2	0.182	0.244
Kids	1.15	1.11	0.080	0.596
≥ High School	0.44	0.46	0.037	0.815
Income ≥ $1,500	0.18	0.25	0.169	0.340
Camp	0.15	0.10	0.159	0.303
"Average" Health	0.36	0.40	0.080	0.610
ISIS Abuse	0.78	0.77	0.018	0.912
Civic Life				
Church ≥ 1x/week	0.58	0.66	0.155	0.322
Assoc. Member	0.30	0.29	0.022	0.887
Had Sunni Friends	0.72	0.70	0.037	0.812
T1 Attitudes				
Agree w/ Gold Adage	0.388	0.33	0.123	0.432
Anti-sectarian	0.70	0.64	0.130	0.409
Coexistence Index	0.07	-0.10	0.178	0.253
Trust Index	-0.07	0.09	0.158	0.325
Iraqi Identity Index	-0.17	0.23	0.406	0.011

ISIS abuse defined as property theft, physical or sexual abuse, torture, arrest, or the kidnappings of family members. "Had Sunni Friends" refers to respondents who report having at least a few Sunni Arab friends before displacement. The gold adage variable refers to agreement with the prejudiced folk saying that "even if a Muslim is a piece of gold, [he] will burn a hole in your pocket." Anti-sectarian refers to agreeing or strongly agreeing with the saying that "Iraq will be a better society if people treat one another as Iraqis, rather than Christians, Shi'is, Sunnis, or Kurds." The Iraqi identity index is imbalanced, but the main results persist when estimated using a difference-in-difference that takes into account the different baseline on this index.

Additional Analyses and Robustness Checks

I present the raw means for each of the six key outcomes across both groups below. Raw means for the three behavioral outcomes align closely although imperfectly with the predicted probabilities for these outcomes (Table 2), which is expected in a small sample size and highlights the importance of adjusting for covariates. The three attitudinal indices are centered at a mean of zero with a standard deviation of 1. Raw means for these indices in both time periods show that treatment effects are driven largely by tolerant shifts among treated participants, and to a lesser extent, negative shifts in trust among control participants.

Table 7: Raw Means (n = 168)

	T1 Control	T1 Treated	T2 Control	T2 Treated
Behavioral Outcomes				
Iftar Attendance	–	–	52.0%	70.0%
			(0.50)	(0.46)
Mixed League Sign-Up	–	–	91.8%	98.6%
			(0.28)	(0.12)
Train w/ Muslims	–	–	28.6%	71.4%
			(0.45)	(0.46)
Attitudinal Outcomes				
Belief in Coexistence	0.05	-0.12	-0.20	0.20
	(1.04)	(0.94)	(0.98)	(0.93)
Iraqi Identity	-0.19	0.21	-0.11	0.08
	(0.94)	(1.03)	(0.87)	(1.16)
Muslim Trust	-0.11	0.02	-0.08	-0.02
	(0.87)	(1.21)	(0.85)	(1.16)

The following series of six plots show the marginal effects of being treated for each key outcome. I do this by generating an empirical distribution of the predicted probabilities for the treatment and control conditions, obtained from 2,000 block-bootstrap simulations of the OLS model. The plots below show the simulated difference between the predicted probabilities for treatment and control units at the 90% confidence level.
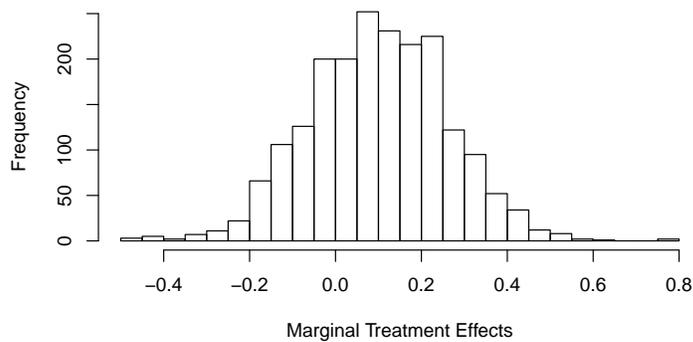
Figure 6: Block-Bootstrapped Treatment Effects (n = 168): Attending *Iftar*
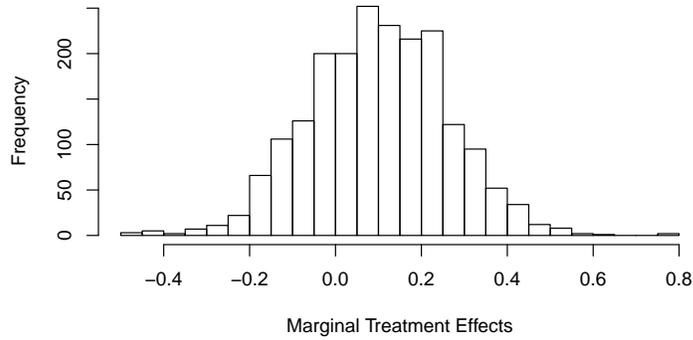
Figure 7: Block-Bootstrapped Treatment Effects (n = 168): Sign-Up for Mixed League
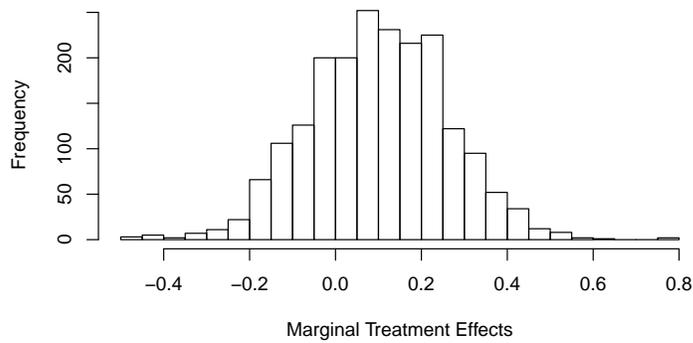
Figure 8: Block-Bootstrapped Treatment Effects (n = 168): Train with Muslims at 4-Months

Figure 9: Block-Bootstrapped Treatment Effects ($n = 168$): Muslim Trust Index

Marginal Treatment Effects
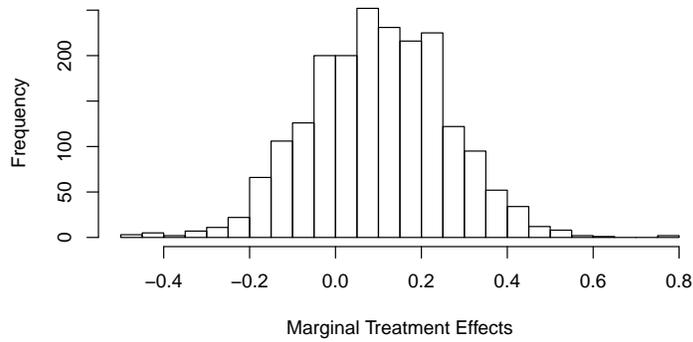
Figure 10: Block-Bootstrapped Treatment Effects ($n = 168$): Belief in Coexistence Index

Marginal Treatment Effects
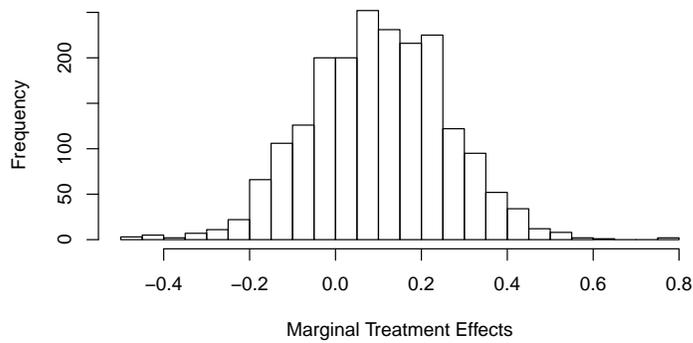
Figure 11: Block-Bootstrapped Treatment Effects ($n = 168$): Iraqi Identity Index
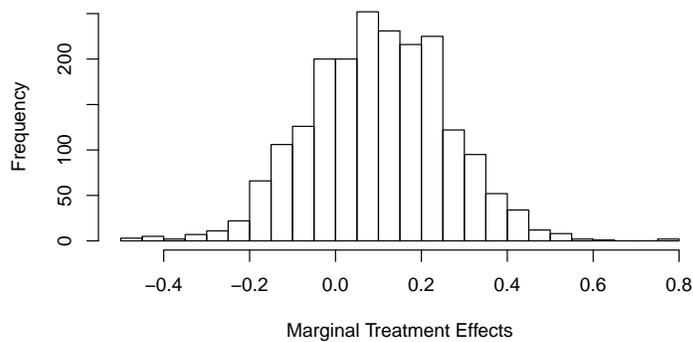

Marginal Treatment Effects

Table 8 below reports OLS results for the six main outcomes with the same models used to generate predicted probabilities in 2.

Table 8: Main Regression Analyses

	Iftar	Mixed League	Muslim Train	Identity	Trust	Coexistence
	Dependent Variables					
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.07	1.10***	0.54	−1.12*	−1.80***	0.76
	(0.32)	(0.24)	(0.43)	(0.57)	(0.56)	(0.56)
Treated	0.26*	0.08*	0.44*	0.20	0.13	0.33**
	(0.15)	(0.04)	(0.26)	(0.17)	(0.17)	(0.16)
Block Bootstrap S.E.	✓	✓	✓	✓	✓	✓
Demographic Controls	✓	✓	✓	✓	✓	✓
Muslim Prejudice Controls (T1)	✓	✓	✓	✓	✓	✓
Observations	168	168	167	168	168	168
R^2	–	–	–	0.08	0.09	0.09
Adjusted R^2	–	–	–	0.01	0.03	0.03

Note: *p<0.1; **p<0.05; ***p<0.01

Demographic controls include age, number of children, whether one lives in a camp, education, home ownership, pre-displacement income, church attendance, abuse experienced by ISIS, and whether the player was an added or core team member. T1 Muslim prejudice controls are: agreement with the derogatory adage that "even if a Muslim is a piece of gold, [he] will burn a hole in your pocket", how welcoming one finds Sunni Arabs, and how one characterizes the relationship between Christians and Muslims before the 2003 war. Only this final prejudice variable is included for models 3 – 5, where the outcome index includes items co-linear with T1 trust items. All variables coded in a pro-trust direction. Missing data imputed using medians. Standard errors clustered at the team level.

Table 9: Permutation Test

	Treat Coef.	p-value
Behavioral Outcomes		
Iftar Attendance	0.26	0.000
Mixed League Sign-Up	0.08	0.021
Train w/ Muslims	0.44	0.000
Attitudinal Outcomes		
Belief in Coexistence	0.32	0.029
Iraqi Identity	0.26	0.214
Muslim Trust	0.16	0.448

Permutation test based on 5,000 samples. Entries are predicted values based on an OLS model with demographic controls for age, number of children, whether one lives in a camp, education, home ownership, pre-displacement income, church attendance, abuse experienced by ISIS, and whether the player was an added or core team member. T1 Muslim prejudice controls are: agreement with the derogatory adage that "even if a Muslim is a piece of gold, [he] will burn a hole in your pocket", how welcoming one finds Sunni Arabs, and how one characterizes the relationship between Christians and Muslims before the 2003 war. Only this final prejudice variable is included for for the attitudinal outcomes, as items in these indices are colinear with T1 trust items. All variables coded in a pro-trust direction. Missing data imputed using medians.

I use an unsupervised hierarchical model to detect latent clusters among the survey items. This data-driven method of identifying latent dimensions generates the scree plot below, indicating that there are three clear latent variables present in the data. The items in these indices align closely with theoretic expectations and can be labeled as trust toward Muslims, belief in coexistence, and the salience of Iraqi identity (Table 1).

Figure 12: Scree Plot of Survey Item Clusters
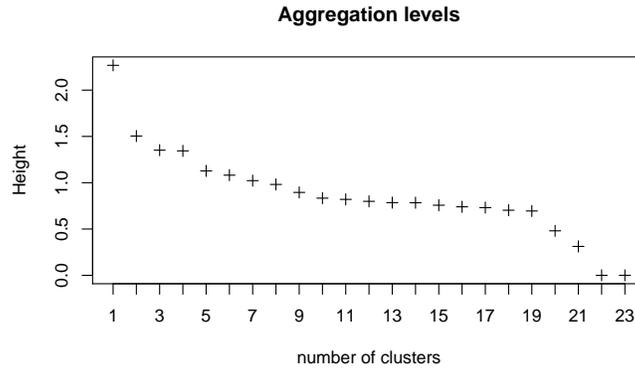
The scree plot generated by the clustering algorithm points to the presence of three clearly delineated latent dimensions in the survey data, and three less clearly defined but nonetheless evident clusters. I take the latter three clusters as secondary attitudinal outcomes, which can be labeled as trust toward non-Christians, freedom and security, and optimism regarding the future. Regression analyses on these outcomes do not yield statistically significant treatment effects, however, perhaps because contact with Muslims is not particularly relevant to these dimensions (11).

Table 10: Secondary Attitudinal Outcomes

Attitudinal Indices

1. Trust toward non-Christians

    1.1 Would be comfortable with a non-Christian as a neighbor
    1.2 Would be comfortable with a non-Christian as a co-worker
    1.3 Would be comfortable with a non-Christian as a classmate for their child

2. Freedom and Security

    2.1 Would move back to hometown if it were rebuilt
    2.2 Can freely practice religion
    2.3 Believe that democracy is suitable for Iraq

3. Optimism

    3.1 Believe that life in 5 years will be more safe than now
    3.2 Believe that life in 5 years will be more happy than now

---

Table 11: Secondary Attitudinal Outcomes

	Dependent Variables		
	Optimism	Security	General Trust
	(1)	(2)	(3)
Intercept	−1.58***	−1.32**	−0.47
	(0.82)	(0.74)	(0.40)
Treated	−0.08	0.06	−0.03
	(0.24)	(0.22)	(0.24)
Team Clustered S.E.	✔	✔	✔
Demographic Controls	✔	✔	✔
Observations	164	168	168
Clusters	14	14	14
R^2	0.17	0.14	0.03
Adjusted R^2	0.10	0.08	−0.03

Note: *p<0.1; **p<0.05; ***p<0.01

The contact literature predicts that both naturalistic intergroup exposure and contact in conflict zones can lead to negative effects on outcomes, stemming from misunderstandings or even violence (Paluck, Green and Green 2017). I test this "adversarial contact" hypothesis by measuring whether red and yellow cards increase when control participants face Muslim opponents, finding that there is no effect of adversarial contact on match-level aggression. I present qualitative evidence in favor of an in-group policing solution to the potential for adversarial contact to foment violence. To estimate in-group policing, I leverage the random assignment of a Muslim vs. Christian referee to each game. Table 12 shows that Muslim referees are systematically associated with less aggressive matches, even when controlling for adversarial contact (captured by mixed Muslim-Christian vs. all-Christian opponents).

Table 12: Referee Identity and Infractions

	Dependent Variables:		
	Total Cards	Yellow	Red
	(1)	(2)	(3)
Intercept	4.46***	4.01***	0.45
	(1.05)	(0.83)	(0.34)
Mixed	0.47	0.46	0.01
	(0.72)	(0.57)	(0.23)
Both Treated	0.29	0.17	0.12
	(0.86)	(0.68)	(0.28)
Muslim Ref	−3.74*	−3.26**	−0.48
	(2.03)	(1.60)	(0.65)
Team Clustered S.E.	✓	✓	✓
Goal Controls	✓	✓	✓
Referee F.E.	✓	✓	✓
Observations	98	98	98
R^2	0.17	0.24	0.04
Adjusted R^2	0.01	0.09	−0.14

Note: $^*p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

While this study has focused on improving Christian trust toward Muslims, the intolerance of Sunni Arabs toward Christians fuels the sectarian conflict in Iraq and bears investigation. A sample size of 28 Muslims is too under-powered to detect over-time changes in tolerance toward Christians, but the means presented in Table 13 are suggestive of positive effects of contact among Muslim participants as the league progressed.

Table 13: Changes in Muslim Prejudice (Difference in Means Tests)

Survey Item	T1	T2	p-value
Anti-sectarian (1)	2.79	2.75	0.867
Anti-sectarian (2)	1.46	1.79	0.323
General trust	0.15	0.25	0.353
Happiness	0.42	0.54	0.417
Perceived Safety	0.42	0.50	0.579
Tolerance Important	0.96	0.96	1
Proud Iraqi	2.39	2.43	0.887

Difference in means t-tests applied to data sub-set by baseline (T1) and post-treatment (T2) responses ($n = 28$).

In line with the hypotheses outlined in the "What Types of Civic Groups Can Build Trust?" section, the effects of intergroup contact may vary based on: prior contact with Muslims, pre-treatment prejudice toward Muslims, and membership in civic organizations. Subsetting the data down to groups with low/high contact and low/high prejudice shows and re-running the OLS models shows that those who report having Sunni Arab friends and visiting Sunni Arab shops regularly before displacement are associated with higher treatment effects relative to those who do not report these levels of contact. Conversely, those with Sunni Arab neighbors are associated with lower treatment effects, possibly because of the widespread (and well-documented) belief that Sunni Arabs colluded with ISIS to identify Christian houses. Under duress or not, this collaboration was seen as a betrayal by many Christians (Hall 2017). The sample size is not sufficiently powered to handle treatment effects rather than splitting the data in this way, except for one interactive model that yielded statistically significant interaction terms (Table 14, Model 6). This model suggests that being a member of a civic organization significantly amplifies treatment effects, suggesting that the skills and norms learned in one civic context can translate and intensify when re-enforced in others.

Table 14: Heterogenous Effects by Prior Contact, Baseline Prejudice, and Civic Membership

	Dependent Variables								
	Friends		Shops		Neighbors		Gold Adage		Association
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Constant	0.20	−0.15	0.52*	−0.33	0.66*	−0.20	0.01	−0.02	0.82
	(0.30)	(0.50)	(0.27)	(0.40)	(0.39)	(0.44)	(0.58)	(0.19)	(0.46)
Treated	0.31**	0.15	0.27*	0.15	0.29**	0.15	0.24	0.46*	0.20
	(0.15)	(0.16)	(0.16)	(0.15)	(0.11)	(0.18)	(0.18)	(0.17)	(0.19)
Assoc : Treat	–	–	–	–	–	–	–	–	0.46*
									(0.28)
Team Clustered S.E.	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	101	67	88	77	64	100	61	107	168
R^2	0.29	0.15	0.28	0.28	0.33	0.19	0.15	0.21	0.11
Adjusted R^2	0.21	−0.02	0.18	0.16	0.19	0.09	−0.06	0.11	0.02

Note: *p<0.1; **p<0.05; ***p<0.01

Models are the same as those used for the main analyses (Table 8) and subsetted to relevant populations. Odd-numbered models refer to analyses subsetted to those with low contact (low tolerance for the gold adage analysis in models 7 and 8) while the even-numbered models are subsetted to those with high contact or tolerance. This bifurcation of the data for the models 1 – 6 is based on a binary distinction between low and high ("many" or "few" vs. "never"). The subset for model 7 is based on rejecting the prejudiced folk saying that "even if a Muslim is a piece of gold, [he] will burn a hole in your pocket" and model 8 analyzes the subset that agreed with this saying. Model 9 is run on the full sample and includes an interaction term, the only such model with a statistically significant interaction effect given the small sample size.

Table 15: Statistically Significant Individual Survey Items

	Dependent Variables			
	Sunnis Pro-ISIS (1)	Anti-sectarian (2)	Feel Safe (3)	Proud Iraqi (4)
Intercept	1.97***	3.04***	2.25***	3.64***
	(0.25)	(0.62)	(0.53)	(0.50)
Treated	−0.16**	0.28*	0.29*	−0.38**
	(0.07)	(0.17)	(0.15)	(0.18)
Team Clustered S.E.	✔	✔	✔	✔
Observations	168	163	167	168
R^2	0.09	0.09	0.06	0.13
Adjusted R^2	0.02	0.02	−0.01	0.07

Note: *p<0.1; **p<0.05; ***p<0.01