



MIDDLE EAST AND
NORTH AFRICA

REGIONAL PROGRAM WORKING PAPER 24

OCTOBER 2019

Impacts on Trust and Social Capital of a Youth Employment Program in Yemen

Evaluation of the Rural and Urban Advocates Working for Development intervention for the Social Fund for Development

Olivia Bertelli, Sikandra Kurdi, Mai Mahmoud, Mohamad Al-Maweri, and Tareq Al Bass

CONTENTS

Abstract	iv
1. Introduction	1
1.1 The Rural and Urban Advocates Working for Development intervention	1
1.2 Theory of Change	2
2. Data and Methodology	2
2.1 Data	2
2.2 Survey Methodology	3
2.3 Trust Games Methodology	4
3. Survey Results	6
3.1 Changes in Aspirations and Optimism	6
3.2 Changes in Self-Reported Trust and Social Capital	9
3.3 Changes in Self-Efficacy and Views on Gender	13
4. Trust Game Results	14
4.1 Correlation Between Survey Measures and Trust Game Measures of Cooperation	14
4.2 Baseline Cooperativeness by Group	14
4.3 Changes in Measured Cooperativeness	15
5. Conclusions	16
Appendix: Attrition Analysis	18
Attrition in Survey Participation	18
Attrition in Game Participation	20
References	23

TABLES

Table 2.1.1 Geographic distribution of the sample of RUAWFD participants	3
Table 2.3.1 Payoff matrix for trust games for each player, in Yemeni riyals	5
Table 3.1.1 Own aspirations for permanent job	7
Table 3.1.2 Aspirations for building or buying a house	7
Table 3.1.3 Aspirations for buying a car	8
Table 3.1.4 Perceptions of job market demand for the respondent's skills and specialization	8
Table 3.1.5 Optimism about the future	8
Table 3.2.1 Trust in people	9
Table 3.2.2 Trust in institutions	10
Table 3.2.3 Perceptions of people in the community	11
Table 3.2.4 Knowledge about cooperation in community	11
Table 3.2.5 Youth decision making in your community	12
Table 3.2.6 Issues around which young people can participate in decision making	12
Table 3.2.7 Perceptions and aspirations for community cooperation	12
Table 3.3.1 Average responses to questions on gender equality	13
Table 3.3.1 Self-efficacy	14
Table 4.1.1 Correlations between survey measures and trust game measures	14
Table 4.2.1 Baseline cooperativeness by group: main sample	15
Table 4.3.1 Changes between baseline and follow-up in main sample	16

Appendix Table 1 Sample size by survey rounds.....	18
Appendix Table 2 Attrition analysis for the survey panel	19
Appendix Table 3 Sample size for games	20
Appendix Table 4 Balance test for participants in first vs. second round of games.....	22

FIGURES

Figure 2.3.1 Governorate groupings for partner pairing.....	6
---	---

ABSTRACT

This paper evaluates the impacts on the participants of the Yemen Social Fund for Development's youth employment and training program called Rural and Urban Advocates Working for Development (RUAWFD). The evaluation used both traditional surveys and an innovative experimental game methodology to show that the employment program, in addition to aiding youth individually, has important benefits for the country as a whole by contributing to stronger social capital.

The survey analysis finds for the program participants significant increases between the baseline and follow-up surveys in self-reported trust in local government institutions and officials, political parties, and tribes. In reflecting on the level of cooperativeness in their own communities, participants reported increased awareness of the presence of marginalized groups and increased perception of cooperativeness in surrounding communities. There was also a significant increase in self-reported trust in people generally, especially for trust in other young people and in people from other areas of Yemen.

The experimental game methodology uses a common pool game from the experimental economics literature incentivized by cash payments to measure trust levels between pairs of RUAWFD participants from different geographic regions. This approach confirms the findings from the survey analysis while avoiding possible self-reporting bias. The game results show that trust was lowest at baseline for partners in which one of the partners was from one of the Northern governorates and the other was from one of the Southern governorates. After the intervention, however, not only were average trust levels higher, but Northern-Southern pairs of RUAWFD participants had trust levels closer to those for pairs from the same regions.

These findings are consistent with the literature on inter-group contact theory suggesting that community interventions can increase trust in individuals and institutions. This research contributes to a growing literature on trust and social capital as important development indicators, particularly in relation to conflict. The main results suggest that reinforcing social ties across regions in Yemen is an important benefit of the Social Fund for Development's role as a national development agency and an achievable objective to consider in planning development interventions to contribute to future post-conflict reconstruction.

1. INTRODUCTION

The conflict in Yemen, which started in 2015, has been incredibly destructive in terms of lives lost, physical infrastructure destroyed, economic setbacks, and deterioration of human development indicators. Looking towards the future, a long-term fear is that the conflict has also led to a decrease in intergroup trust. This breakdown in bridging social capital as a result of conflict is theorized to be part of the conflict trap that can prevent reconstruction and rebuilding of a stable society (Collier et al. 2004). The Yemen Social Fund for Development (SFD) is a unique national institution that works across all governorates in Yemen, including those controlled by warring parties in the current civil war. The Rural and Urban Advocates Working for Development (RUAWFD) intervention is a youth employment and training program which was launched by SFD in order to increase the capacity of unemployed young people while also supporting development projects in poor communities. This intervention also has promise as a way to create cross-cutting ties between participants from different parts of a currently fractured country.

1.1 The Rural and Urban Advocates Working for Development intervention

The RUAWFD intervention is both a capacity building program targeting university graduates and a mechanism for employing university graduates to promote community-led development projects. In this evaluation, we focus on the impacts of the program for the youth participants along three dimensions identified as general program objectives: 1) capacity building and preparation for the labor market; 2) enhancing development awareness; and 3) reinforcing social capital.

The selection criteria for participating in the program include having a university degree and being under the age of 35 at the start of the program. In addition, all program participants were required to attend a training led by SFD on rural development before joining RUAWFD. The selection of participants also considered gender balance between male and female applicants.

The duration of the program was fifty days, divided between training workshops and field work. Training workshops were done in two rounds of 10 days each, while the field work was done in two rounds of 15 days. The curriculum of the workshops was categorized into three parts: a) background knowledge on community work; b) development issues and requirements; and c) teamwork and communication skills. In addition, the training was customized to meet the specific needs of each governorate.

The initiatives implemented by RUAWFD participants during their field work in rural communities included activating village cooperation councils; providing relief to internally displaced families; building rural roads; and raising awareness of issues related to health, female education, and women participation in community initiatives. In total, RUAWFD covered 26 districts across 13 governorates.

The program training and field work were directly designed to build the capacity of youth to prepare them for the labor market and to enhance their development awareness. We measure these outcomes in terms of the perception of the participants on their own labor market readiness and on the knowledge and skills that they acquired related to development projects that can be implemented in their communities. The dimension of enhancing social capital is less obvious, but implicit in the design of the program. While the youth from different regions who participated in the program did not interact in person, they received similar training, including a human-rights based approach to development, had similar roles, goals, and experiences in the field, and, through a WhatsApp group, are part of a virtual network of all who have participated in the RUAWFD program.

1.2 Theory of Change

Recent literature in economics has identified social capital as a key driver of community development. One element of social capital is strong intra-community ties. However, equally important is the bridging type of social capital which creates links that cross social divides based on religion, class, ethnicity, gender, or socio-economic status (Woolcock 2000). This bridging type of social capital is important for avoiding sectarianism and for rebuilding social capital in a post-conflict setting by nurturing cross-cutting ties between opposing groups (Colletta and Cullen 2000). Evidence of breakdown in the cross-cutting ties element of social capital has been found in post-conflict settings through the use of experimental games to measure cooperation and trust. Experiments have found that experiencing hardship, while it increases cooperation with members of the same group, it also decreases trust in members of opposing groups (Gilligan, Pasquale, and Samii 2014; Bauer et al. 2014).

Attempts at rebuilding social capital in post-conflict settings have found that even brief exposure to new institutions with cross-cutting participation can increase inter-group cooperation (Fearon et al. 2009). Experiments testing intergroup contact theory in different settings have found that contact between opposing groups is more successful at reducing prejudice when the experience occurs in a program, such as RUAWFD, that provides participants with common goals, emphasizes a cooperative environment, brings together groups with equal status, and demonstrates authority sanction for the contact (Pettigrew and Tropp 2006).

This background suggests that RUAWFD has the potential to create bridging social capital by connecting young people in different areas of the country in a common development initiative.

2. DATA AND METHODOLOGY

2.1 Data

Our analysis focuses on participants in RUAWFD projects which started between October 2017 and March 2018. All participants in these projects were requested to complete a baseline survey questionnaire in November 2017 as part of their initial training session, and a follow-up questionnaire in October 2018. All participants were randomly assigned to play the trust games either prior to their participation in RUAWFD (first round of games in October 2017) or after their participation in RUAWFD (second round of games in October 2018). Among all 1045 youths, some 202 had already taken part in the RUAWFD program in the past (“old participants”), while for the others it was their first time (“new participant”). Table 2.1.1 breaks down the number of new and old participants by governorate.

Some youth signed up for the program and participated in the trust games but dropped out before participating in the baseline survey. Other youth enrolled in the program after the trust game exercise and baseline survey had been completed. Also, one project in Bani Mattar district in Sana’a was cancelled due to rejection by community leaders. These youth are not included in our analysis. The sample we use is based on all participants who filled in the baseline survey, indicating that they participated in the initial training.

Table 2.1.1 Geographic distribution of the sample of RUAWFD participants

Governorate	New participants	Old participants	Total
Ibb	46	45	91
Abyan	103	0	103
Al-Bayda	18	21	39
Hudaydeya	26	0	26
Al-Dhalea	111	0	111
Maharah	45	0	45
Taiz	136	0	136
Hajjah	93	37	130
Hadhramout	56	0	56
Dhamar	58	29	87
Sana'a	0	45	45
Amran	76	25	101
Maareb	75	0	75
Total	843	202	1,045

Source: Authors' analysis.

2.2 Survey Methodology

The survey questionnaire included modules on:

- Work experience (baseline only)
- Social network
- Trust in people
- Trust in institutions
- Own aspirations
- Community aspirations
- Optimism about the future
- Conflict exposure (baseline only)
- Self-efficacy
- Social solidarity
- Female empowerment
- Program participation (follow-up only)

The graduates filled in the survey on paper at the training centers, then a process of double entry of the data was carried out by staff at SFD headquarters in Sana'a. While the participants are relatively highly educated and could clarify any unclear questions with program staff who were trained on the survey content and structure, the fact that the survey was self-administered likely contributed to a higher level of missing data through participants omitting to answer some of the questions.¹

We employ a simple first difference strategy to test for changes in participant characteristics between the baseline and follow-up surveys. Although this approach is not rigorously causal, as it does not allow us to distinguish impacts of the program from changes that would have occurred over time in the absence of the program, it does provide suggestive evidence. We believe that in the current context in Yemen there is little reason to assume any positive time trends in optimism, aspirations, or trust in people in other parts of Yemen in the period between the two surveys. Consequently, to the extent that attitudes changed independently of the program, our results are

¹ The most commonly skipped question was the general question about trust in people. Because the labelling for the Likert index was more detailed for this question, the first question in the series, it may have appeared to some respondents to be an example of the format of the questions in the section rather than a question that they should answer.

likely to be underestimates of the true program impact. The length of time between the two surveys was also fairly short at less than one year at the maximum and generally only a few months. This is because the surveys were administered prior to and after participation in the program for groups of youth participating in the same locally administered project within the overarching RUAWFD program.

Our analysis focuses on youth participating for the first time in the RUAWFD program. Some youth who had previous experience in the program were also surveyed at baseline but were excluded from the analysis because the degree to which they were affected by the experience was expected to differ from new participants. The sample size for resurveyed new participants is 788. Appendix I includes detailed attrition analysis and reasons for changes in sample size between baseline and follow-up surveys.

Our first approach is to estimate changes only among the sample of new participants in the program and for whom we have both baseline and follow-up survey information. For this specification, we include individual fixed effects.

$$y_{it} = \alpha + followup_t + \gamma_i + u_{it} \quad (1)$$

where y_{it} represents the outcome of interest (e.g. self-reported trust, personal aspirations, etc.), $followup_t$ is a dummy variable taking values 0 for the baseline survey results and 1 for the follow-up survey, α is a constant term, γ_i represents individual fixed effects, and u_{it} is the error term, which is clustered at the local project level.²

2.3 Trust Games Methodology

The public good game is widely used as measure of cooperativeness in experimental economics (Johnson and Mislin 2011). The advantage of games compared to surveys at measuring cooperative attitudes is that participants in games have a monetary incentive to act according to their actual beliefs, so the measures of cooperativeness from games are less vulnerable to social desirability bias.

The basic format of the game is as follows:

- Each partner receives an initial amount of money that they can choose to either invest in a group project or save for themselves.
- The total pool of money invested in the group project is increased by 50 percent and then divided in half.
- Each player receives half of the returns from the group project plus any money that they decided to save for themselves.

This setup generates a “Prisoner’s dilemma” situation where the optimal high equilibrium with both players contributing to the public good is only reached if there is sufficient trust between the players.

In our games, the initial amount was 2500 Yemeni riyals (about USD 7). Table 2.3.1 illustrates examples of possible outcomes for selected strategies of the two players. As can be seen, in the absence of trust, the Nash equilibrium is for both players to contribute nothing, as this is the only

² To check whether the results were biased by any form of differential attrition, we alternatively used project-level fixed effects so that all survey data could be included. Covariates were added to control for characteristics that differed at baseline between attriters and non-attriters. The results obtained with the two specifications are very similar. Results for the specification with attrition controls are available upon request.

strategy for which neither player has a profitable deviation. The optimal combined solution for both players if they coordinate their strategies is for both to contribute the full amount 2500.

Table 2.3.1 Payoff matrix for trust games for each player, in Yemeni riyals

Player A → Player B ↓	Contribute 2500	Contribute 2000	Contribute 500	Contribute 0
Contribute 2500	A: 3750 B: 3750	A: 3875 B: 3375	A: 4250 B: 2250	A: 4375 B: 1875
Contribute 2000	A: 3375 B: 3875	A: 3500 B: 3500	A: 3875 B: 2375	A: 4000 B: 2000
Contribute 500	A: 2250 B: 4250	A: 2375 B: 3875	A: 2750 B: 2750	A: 2875 B: 2375
Contribute 0	A: 1875 B: 4375	A: 2000 B: 4000	A: 2375 B: 2875	A: 2500 B: 2500

Source: Authors' compilation.

We use the amount that each player contributes as a measure of their trust in their partner. Partner assignment is explained more fully below.

The main outcome for our analysis is how much more were participants willing to cooperate with a geographically, culturally, or politically similar “near” partner compared to a more distant “far” partner. In particular, the “far” partner is someone who is geographically associated with the opposite side in the current civil conflict. Our measure of cooperativeness is the number of riyals invested in the common pool, so we define an outcome variable *NearPreference* as:

$$NearPreference_i = AmountNear_i - AmountFar_i \quad (2)$$

Where *AmountNear* is the amount of riyals invested during the game with a “near” partner and the *AmountFar* is the amount of riyals invested during the game with a “far” partner.

For the trust games, participants who were included in the original sampling frame were randomly assigned to play the trust games either before or after their participation in the RUAWFD program. In each project location, all participants played either exclusively at baseline or exclusively at follow-up to avoid communication between participants about the games during the program implementation contaminating our results.

Each participant was assigned two different partners, one for the “near” game and one for the “far” game. Ideally, the “near” game was with a same-gender participant from the same region and the “far” game was with a same-gender participant from the opposite region. Given that there were far more participants from the North than from the South, participants were divided into three regional grouping rather than two: South, North-Central & Tihama, and North. The governorate of residence of the partner was made salient during the games by originally introducing the two partners to each other with a text message which shared their partners’ name as well as governorate and phone number.

Limiting “near” partners to being from the same region but not the same project location ensured that, just like the “far” partners, they would not have met in person. The intended three categories of “far” grouping were:

- Southern governorates with Northern governorates
- Southern governorates with North-Central & Tihama governorates
- Northern governorates with North-Central & Tihama governorates

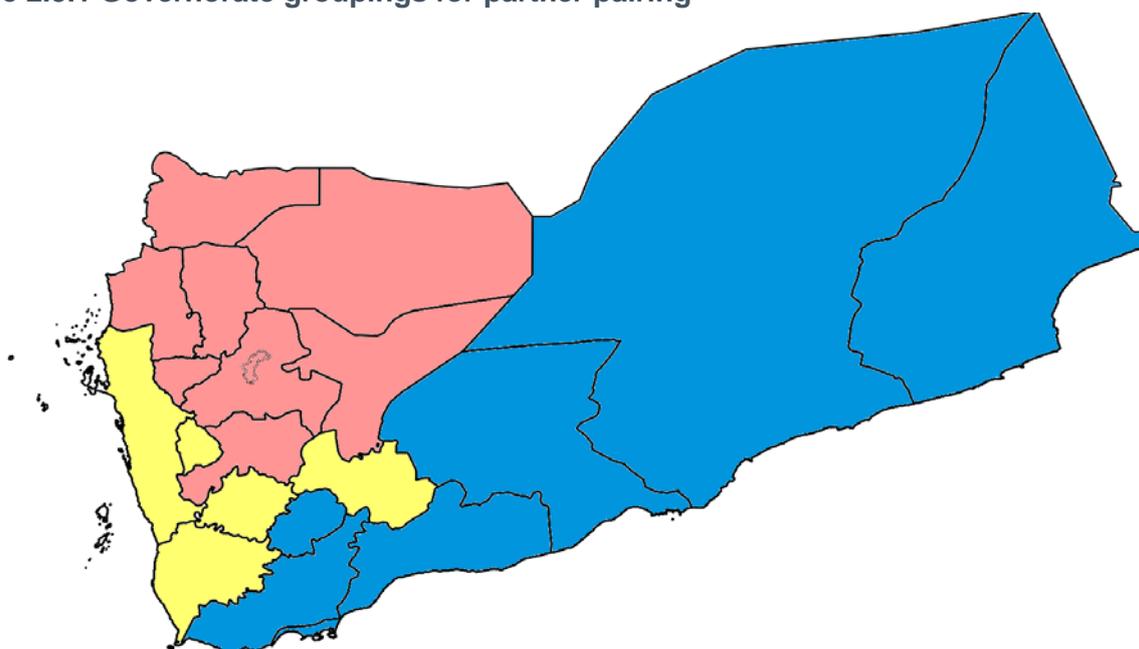
While the “near” groupings were:

- Southern governorates with Southern governorates (but not the same project location)
- North-Central & Tihama governorates with North-Central & Tihama governorates (but not the same project location)
- Northern governorates with Northern governorates (but not the same project location)

To avoid learning effects being correlated with partner location, the order in which the near and far games were played was also randomized.

As described below, some issues with imperfect randomization resulted in not all participants being assigned to one of these intended groupings.

Figure 2.3.1 Governorate groupings for partner pairing



Code	Legend	Group	Governorates in group
S	Blue	Southern Governorates	Abyan, Aden, Al-Dhalea, Hadhramout, Lahj, Maharah
N1	Yellow	North-Central & Tihama	Al-Baydah, Hodeidah, Ibb, Raimah, Taiz
N2	Red	Northern Governorates	Al-Jawf, Amran, Dhamar, Hajjah, Maareb, Mahweet, Saadah, Sana'a

Source: Authors' compilation.

As in the survey data, there was substantial attrition between the planned sample and the final sample. Due to uneven group sizes, some participants were unable to be properly matched with both a “near” and “far” partner. A further reduction in sample size was caused by a mistake in coding some of the governorate locations where the text field was left blank. Our final sample size using the three randomly assigned groups is 258 participants at baseline and 324 at follow-up. This is the sample size used when using the outcome variable “Near Preference” and showing the difference in outcome by group assignment. We explain our checks for attrition bias in Appendix I.

3. SURVEY RESULTS

3.1 Changes in Aspirations and Optimism

Between the baseline and follow-up survey, we find significant increases for new participants in their perception of the likelihood of attaining a permanent job, house, or car and in their perception of the demand for their skills on the labor market. However, we do not find any significant changes in

measures of overall optimism about personal situations of the respondents or the general situation of Yemen.

The questions on aspirations capture the respondents' ambitions and optimism about the future. In terms of ambitions, respondents were asked whether they wanted to have a permanent job, own their own house, and own a car. For the vast majority of respondents who indicated they do desire to reach these goals, there was a set of two follow-up questions asking the respondent about the likelihood – on a scale from 1 (low likelihood) to 10 (high likelihood) – of achieving this goal both under normal (i.e., absent the conflict) conditions and under current conditions. Results reported in

Table 3.1.1 does not show any significant change in the share of respondents who want to have a permanent job, as the share was already at 98 percent. This result assures us that the changes in perceived likelihood are not biased by any change in the share of respondents answering this introductory question. There also was no significant change in respondents' evaluation of the likelihood of obtaining a permanent job under current conditions. However, we find a significant increase in the perceived likelihood of attaining a permanent job under normal conditions in the range of 0.27 points on a 10-point scale.

Table 3.1.1 Own aspirations for permanent job

	(1) Desire Permanent Employment	(2) Likelihood (1-10) Currently	(3) Likelihood (1-10) Normal
Follow-up	0.001 (0.857)	0.066 (0.734)	0.273** (0.020)
Observations	1,567	1,476	1,481
Mean of dependent variable	0.985	5.297	7.450

Source: Authors' analysis. Note: p-values in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01. Variables (2) and (3) are based on a ten-point scale ranging from 1 (low likelihood) to 10 (high likelihood).

Table 3.1.2 shows regression results for the questions on respondents' aspirations for owning their own house. There is a significant increase in the aspirations for owning a house by around 4 percent. Given this positive change in answers to the introductory question on house ownership aspirations, the follow-up questions on the likelihood of owning a house under current and normal conditions have also significantly increased by around 0.30 and 0.50 points, respectively.

Table 3.1.2 Aspirations for building or buying a house

	(1) Build or Buy House	(2) Likelihood (1-10) Currently	(3) Likelihood (1-10) Normal
Follow-up	0.036*** (0.001)	0.317** (0.030)	0.531*** (0.000)
Observations	1,564	1,468	1,475
Mean of dependent variable	0.955	3.572	5.779

Source: Authors' analysis. Note: p-values in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01. Variables (2) and (3) are based on a ten-point scale ranging from 1 (low likelihood) to 10 (high likelihood).

Table 3.1.3 shows regression results for the questions on respondents' aspirations for owning a car. There is a slight increase in the high share of respondents who expressed desire for owning a car, but it is not statistically significant. The respondents' perception of the likelihood of attaining this goal under both current and normal conditions has significantly increased by around 0.50 and 0.60 points, respectively.

Table 3.1.3 Aspirations for buying a car

	(1) Buy Car	(2) Likelihood (1-10) Currently	(3) Likelihood (1-10) Normal
Follow-up	0.019 (0.142)	0.468*** (0.000)	0.574*** (0.000)
Observations	1,563	1,459	1,460
Mean of dependent variable	0.939	3.109	5.499

Source: Authors' analysis. Note: p-values in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01. Variables (2) and (3) are based on a ten-point scale ranging from 1 (low likelihood) to 10 (high likelihood).

In addition, respondents were asked about their perceptions of job market demand for their skills and specialization under current versus stable conditions. The phrasing used was similar to "what do you think is the probability of demand for your skill compared to others in the labor market in current conditions". The answers were on a scale from 1 (low likelihood) to 10 (high likelihood).

Results in Table 3.1.4 indicate a significant increase in perceptions of job market demand for the respondents' specialization, both under current conditions (by a magnitude of 0.30) and under normal conditions (by a magnitude of 0.40). Likewise, the perceptions of job market demand for the respondents' skills has significantly increased by about 0.50 points for both the question on current conditions and on normal conditions.

Table 3.1.4 Perceptions of job market demand for the respondent's skills and specialization

	(1) Specialization Current Demand	(2) Specialization Normal Demand	(3) Skills Current Demand	(4) Skills Normal Demand
Follow-up	0.333** (0.046)	0.390*** (0.001)	0.507*** (0.001)	0.452*** (0.000)
Observations	1,546	1,543	1,538	1,527
Mean of dependent variable	5.653	7.307	6.257	7.783

Source: Authors' analysis. Note: p-values in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01. Variables are based on a ten-point scale ranging from 1 (low likelihood) to 10 (high likelihood).

Optimism about the future was captured by three questions asking respondents about the expected future situation for the country as whole, for their family's economic situation, and for themselves individually. The answer choices for the questions on optimism consisted of 5 levels: 1 (much better), 2 (somewhat better), 3 (same situation), 4 (somewhat worse), and 5 (much worse).

Table 3.1.5 shows that the average respondent is fairly optimistic about the future. The average response to the three questions on the future situation in Yemen, the future economic situation of the respondent's family, and the respondent's personal situation in the future was just below 2 (somewhat better). We do not find any significant differences in these measures between the baseline and follow-up surveys.

Table 3.1.5 Optimism about the future

	(1) Situation in Yemen	(2) Family economic situation	(3) Personal situation
Follow-up	-0.054 (0.224)	-0.009 (0.820)	0.035 (0.206)
Observations	1,572	1,571	1,569
Mean of dependent variable	1.923	1.811	1.569

Source: Authors' analysis. Note: p-values in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01. Variables are based on a five-point scale ranging from 1 (much better) to 5 (much worse).

3.2 Changes in Self-Reported Trust and Social Capital

We find significant increases in self-reported trust in all categories of people outside of the respondent's immediate family, as well as in local government institutions and government officials, political parties, and tribes. In reflecting on the level of cooperativeness in their own communities, participants reported increased awareness of the presence of marginalized groups and increased perception of cooperativeness in surrounding communities. Both of these results may have colored the mixed results for changes in perceived cooperativeness in the participants' home community.

Trust in people

Self-reported trust was measured by a set of questions asking the respondent to rank their trust in several groups of people as well as various institutions on a scale from 1 (no trust) to 5 (complete trust), following standard questions used in the World Values Survey. The baseline level of trust in our sample at 39 percent is extremely close to the one measured by the 2014 World Value Survey, 38.5 percent. The series of questions on trust in people started with a standard question about the respondent's trust in people in general: "Generally can you say that most people are trustworthy (you do not have to be too careful in dealing with them)." This question was then followed by a set of questions on the perceived trustworthiness of family members, people in the respondent's village or neighborhood, people in the respondent's district, people in the respondent's governorate, people in other regions of Yemen, the older generation in general, and the younger generation.

Table 3.2.1 shows the regression results for the questions on trust in people. There are significant increases in self-reported trust in all groups of people considered. Trust in family members is the only category that had a small but statistically significant decrease between the baseline and follow-up surveys. However, this is likely explained by censoring at the maximum of our scale due to the very high level of reported trust in family at baseline. We show below that these self-reported trust levels also correlate well with the cooperativeness outcomes measured using the trust games.

Table 3.2.1 Trust in people

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	General	Family members	People in your neighborhood	People in your district	People in other governorates	People in other regions of Yemen	Older generation	Younger generation
Follow-up	0.230*** (0.000)	-0.120*** (0.001)	0.198*** (0.000)	0.324*** (0.000)	0.334*** (0.000)	0.338*** (0.000)	0.148** (0.013)	0.197** (0.010)
Observations	1,293	1,572	1,565	1,563	1,560	1,553	1,564	1,564
Mean dependent variable	3.480	4.734	3.662	3.118	2.858	2.636	3.985	3.285

Source: Authors' analysis. Note: p-values in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01. Variables are based on a five-point scale ranging from 1 (no trust) to 5 (complete trust).

Trust in institutions

The questions on trust in institutions focused on the respondent's trust in SFD, tribes, local government, non-government organizations (NGOs), religious leaders, news media, government officials, sheikhs (traditional local leaders), and political parties. During their fieldwork as community organizers, participants interacted directly with leaders of village-level institutions, including sheikhs and tribal and religious leaders. Participants are also expected to have become more aware of perceptions in the community of the responsiveness of local government, political parties, and government officials.

Table 3.2.2 shows regression results for the questions on trust in institutions. There is a significant increase in reported trust in tribes, local government, government officials, and political parties. The institutions for which we find a significant increase in trust between the baseline and endline surveys are those which had the lowest level of trust at baseline.

Table 3.2.2 Trust in institutions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	SFD	Tribes	Local government	NGOs	Religious leaders	News media	Government officials	Sheikhs	Political parties
Follow-up	-0.041** (0.020)	0.206*** (0.002)	0.215*** (0.004)	0.101 (0.108)	-0.044 (0.402)	0.115 (0.124)	0.202*** (0.003)	0.046 (0.445)	0.162*** (0.006)
Observations	1,569	1,566	1,561	1,564	1,564	1,564	1,562	1,568	1,567
Mean dependent variable	4.785	3.356	2.816	3.307	3.593	2.341	2.225	2.686	1.785

Source: Authors' analysis. Note: p-values in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01. Variables are based on a five-point scale ranging from 1 (no trust) to 5 (complete trust).

The only institution with a low level of baseline trust where there was no impact between the baseline and endline surveys was the news media, likely because the program did not change participants' exposure to media. There was also no significant change in trust in religious leaders or NGOs, but this may be because baseline trust in these institutions was already relatively high.

However, the results show a slight decrease in the reported trust in SFD during between the baseline and the follow-up survey. Similarly to the lack of increase found for trust in family members, this is likely due to baseline level of trust in SFD is already almost at the maximum. The magnitude of the change, though statistically significant, is very small (2.3 percent of the sample mean), so should not be seen as discouraging.

Community solidarity and awareness of development issues

While the RUAWFD program was specifically designed to increase social solidarity in the targeted communities, as the focus of this evaluation is on the impact on the youth participants themselves, we wanted to see the degree to which participation in the program changed youths' perception of their role in their home community.

Respondents were asked about the degree to which they agree with the statement regarding their own communities: "People look out mainly for the welfare of their own families and are not much concerned with community welfare" on a scale from 1 (strongly agree) to 4 (strongly disagree). Table 3.2.3 shows regression results for this question and "Do people in your community contribute time and money towards development goals" with answers on a scale from 1 (contribute a lot) to 4 (do not contribute at all). There is a significant decrease in the respondents' perceptions of people's contribution towards common development goals (with a mean decrease of 0.16 on the four-point scale). This decrease may be due to underlying negative time trends. Alternatively, participants having positive experiences during their RUAWFD participation might had higher expectations for their community after seeing higher levels of cooperation in other communities, as is suggested later in the regression analysis results in Table 3.2.7.

Table 3.2.3 Perceptions of people in the community

	(1) Concerned with the welfare of own families rather than community welfare	(2) Contribute to common development goals
Follow-up	-0.014 (0.767)	-0.160*** (0.000)
Observations	1,547	1,564
Mean of dependent variable	2.244	2.162

Source: Authors' analysis. Note: p-values in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01. Variables are based on a four-point scale ranging from 1 (strongly agree) to 4 (strongly disagree) for variable (1) and from 1 (contribute a lot) to 4 (do not contribute at all) for variable (2).

Respondents were also asked a series of yes, no, or 'do not know' questions about the presence of cooperative organizations within the respondent's community through which people volunteer or contribute to the community needs; the presence of development projects that can be implemented in the respondent's community (village or neighborhood); and the presence of marginal groups that are not included in the management of the community or are not engaged in community activities. The purpose of asking these questions in the baseline and follow-up surveys was to test whether the training provided by the program, as well as the experience of working directly on village community development programs, increased participant awareness of cooperative organizations and participation in their own community. In the analysis, we focus on the share of respondents who marked "Do not know" in responses to the questions on marginal groups and cooperative organizations and "No" or "Do not know" in response to the question on the potential for development projects.

Table 3.2.4 shows the regression results for the questions on knowledge about the presence of cooperative organizations, development projects, and marginal groups in the respondent's community. The share of respondents who were unaware of the presence of marginal groups significantly decreased by about 4 percentage points. This result is an indication of the success of the human-rights-based approach of the training materials for the program in helping participants become more aware of what constitutes a marginalized group. The share of respondents who did not have any ideas about development projects in their community also declined by about 4 percentage points.

Table 3.2.4 Knowledge about cooperation in community

	(1) Do not know about the presence of cooperative organizations	(2) Do not know about the potential for development projects	(3) Do not know about the presence of marginal groups
Follow-up	-0.008 (0.724)	-0.042* (0.055)	-0.042** (0.033)
Observations	1,561	1,552	1,547
Mean of dependent variable	0.097	0.085	0.098

Source: Authors' analysis. Note: p-values in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01.

Finally, respondents were asked about youth involvement in community decision-making. Table 3.2.5 shows that the vast majority (95 percent) of the survey respondents find it possible for young people to participate in decision making in their communities. There was no significant change in this measure between the baseline and follow-up surveys. However, when asked a follow-up multiple choice question about the issues in their communities around which young people can participate in decision making (Table 3.2.6), respondents were relatively more likely to choose

"promote community activities and initiatives" in the follow-up survey, which makes sense in line with their participation in the RUWAFD program.

Table 3.2.5 Youth decision making in your community

	(1) Young people participation in decision making
Follow-up	0.007 (0.418)
Observations	1,496
Mean of dependent variable	0.952

Source: Authors' analysis. Note: p-values in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01.

Table 3.2.6 Issues around which young people can participate in decision making

Issue	Baseline	Follow-up	p-value for t-test
Building infrastructure	136 (16.1%)	129 (16.4%)	0.90
Provide support to families in need	363 (43.1%)	311 (39.5%)	0.14
Promote community activities and initiatives	475 (56.3%)	511 (64.8%)	<0.001***
Provide economic support for a particular group or category	84 (10.0%)	72 (9.1%)	0.57
Observations	843	788	

Source: Authors' analysis. Note: frequency in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01

An additional set of questions focused on the cooperation level in surrounding communities, on the level of cooperation in the respondent's community, and on the desired level of cooperation that the respondents aspire for their own community. Answer choices for these three questions were on a scale from 1 (low) to 10 (high). These questions were followed by a categorical, multi-response question on the main issues about which people cooperate in the respondent's community.

Table 3.2.7 shows that the respondents perceived the level of cooperation in the community to be relatively high. On a scale from 1 to 10, the average response for the questions on cooperation in surrounding communities and cooperation in the respondent's community was 7.1 and 7.5, respectively. The analysis shows there was a slight increase in the respondents' perception of cooperation level in their own communities by 0.16 points, and a significant increase in their perceptions of cooperation in surrounding communities by around 0.34 points. However, the already high level of aspirations for cooperation in their own community did not change significantly between the two rounds of the survey.

Table 3.2.7 Perceptions and aspirations for community cooperation

	(1) Cooperation in surrounding communities	(2) Cooperation in own community	(3) Desired cooperation in own community
Follow-up	0.344*** (0.010)	0.164* (0.092)	-0.013 (0.905)
Observations	1,566	1,564	1,557
Mean of dependent variable	7.103	7.512	8.947

Source: Authors' analysis. Note: p-values in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01. Variables are based on a ten-point scale ranging from 1 (low) to 10 (high).

3.3 Changes in Self-Efficacy and Views on Gender

While we expected to find evidence of increases in self-efficacy and attitudes towards gender equality, we actually found a significant negative change in self-efficacy and no change in opinions on gender.

A standard set of questions on self-efficacy asked the respondent to evaluate the truth of some statements such as "my future is determined by things I do and choices I make" on a scale from 1 (completely true) to 4 (not true). With regards to gender equality, the respondent was asked how much they agree on a scale from 1 (strongly agree) to 4 (strongly disagree) with a series of statements developed specifically for this context. In order to avoid acquiescence bias, the individual questions were purposefully varied in terms of whether respondents with a supportive view of gender equality would be expected to agree or disagree.

Table 3.3.1 shows summary statistics for the original questions on gender equality. At baseline, respondents tend to "strongly agree" that "girls should be encouraged to continue their education through university." On average, the respondents at baseline "agree" that "women and men should be equally involved in community decisions" and that "wives and husbands should have equal voice in decisions about their family". However, the respondents also "agree" that "men by their nature are more effective in leadership positions than women" and that "women should not travel without a male relative." Finally, the average response to the statement "the best life for a woman is to stay at home" shows that the respondents tend to "disagree" with this statement at baseline.

Table 3.3.1 Average responses to questions on gender equality

	(1)	(2)	(3)	(4)	(5)	(6)
	Women and men should be equally involved in community decisions	Men by their nature are more effective in leadership positions than women	Women should not travel without a male relative	Girls should be encouraged to continue their education through university	The best life for a woman is to stay at home	Wives and husbands should have equal voice in decisions about their family
Baseline	1.600 (0.621)	1.967 (0.811)	1.745 (0.870)	1.120 (0.357)	3.248 (0.755)	1.711 (0.725)
Follow-up	1.627 (0.660)	1.991 (0.859)	1.756 (0.931)	1.107 (0.345)	3.262 (0.753)	1.739 (0.727)
Total	1.613 (0.640)	1.979 (0.834)	1.750 (0.900)	1.114 (0.351)	3.255 (0.754)	1.725 (0.726)
Observations	1,615	1,607	1,598	1,615	1,597	1,602

Source: Authors' analysis. Note: Answer choices on a scale from 1 (strongly agree) to 4 (strongly disagree). Mean coefficients; standard deviation in parentheses.

Indices for self-efficacy and female empowerment were constructed using the first principal component from a PCA analysis (after inverting the responses where necessary to make higher values correspond to greater empowerment). Because there were several cases where only one or two of the questions had missing values, these missing values were filled in with the average response of that respondent in the section. (Results were qualitatively similar when questionnaires with any missing values in these sections were excluded).

Table 3.3.1 shows that there was a significant decrease in the self-efficacy index between baseline and follow-up by 0.3 (about 0.2 standard deviations) and no significant change in views on gender equality.

Table 3.3.1 Self-efficacy

	(1) Efficacy index	(2) Female empowerment index
Follow-up	-0.300*** (0.003)	0.001 (0.985)
Observations	1,571	1,567
Mean of dependent variable	-0.002	0.008

Source: Authors' analysis. Note: p-values in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01.

4. TRUST GAME RESULTS

4.1 Correlation Between Survey Measures and Trust Game Measures of Cooperation

Table 4.1.1 shows that self-reported trust from the survey (combined into a general index of trust in people and normalized to range from 0 to 1) is significantly correlated with average cooperativeness as measured in the trust game. The dependent variable is the simple average of the amount that each respondent contributed in Yemeni riyals as a measure of their trust in their partner over the two games played. In particular, self-reported trust in people is significantly correlated with the amount contributed in the game played with a more distant “far” partner more than the amount contributed in the game with the more similar “near” partner with a more similar “near” partner. The index of self-reported trust in institutions is not correlated with the amount played in the games. This result makes sense as the games are primarily measuring trust in the other player. Trust in SFD itself is already very high, so there is little impact of players being unsure that they will receive the game payout.

Table 4.1.1 Correlations between survey measures and trust game measures

	(1) Average amount played	(2) Amount far partner game	(3) Amount far partner game	(4) Amount near partner game	(5) Amount near partner game	(6) Amount near partner game	(7) Near preference
Trust in people	402.4* (209.62)		605.8** (240.22)		199.0 (232.44)		-405.2* (218.44)
Trust in institutions		165.6 (232.30)		234.8 (291.36)		96.4 (238.52)	
Observations	563	561	563	561	563	561	564
Mean dependent variable, Yemeni riyals	1651.5	1651.6	1627.4	1626.1	1675.6	1677.1	49.0
R-squared	0.008	0.001	0.013	0.002	0.001	0.000	0.005

Source: Authors' analysis. Note: The dependent variable is the simple average of the two games played. “Near preference” is defined as the difference between the riyals invested by the respondent during the game with a “near” partner and the amount invested during the game with a “far” partner. Standard deviation in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01.

4.2 Baseline Cooperativeness by Group

Table 4.2.1 shows that at baseline Near Preference is significantly higher for partners from regions associated with opposing sides in the conflict.

Table 4.2.1 Baseline cooperativeness by group: main sample

	(1)	(2)	(3)	(4)	(5)
	Near preference	Near preference	Amount average game	Amount far game	Amount near game
Southern and Northern pairing (S_N2)	529.2*** (177.54)	518.6*** (177.41)	-234.5* (129.27)	-499.1*** (158.76)	30.1 (154.85)
Southern and North-central & Tihama pairing (S_N1)	448.7** (189.12)	436.0** (189.03)	-72.6 (137.70)	-296.9* (169.11)	151.8 (164.95)
Near game first		-166.6 (122.26)			
Own region fixed-effects?	Yes	Yes	Yes	Yes	Yes
Observations	258	258	258	258	258
Mean dependent variable, Yemeni riyals	97.5	97.5	1633.8	1585.1	1682.6
R-squared	0.054	0.061	0.843	0.772	0.799

Source: Authors' analysis. Note: "Near preference" is defined as the difference between the riyals invested by the respondent during the game with a "near" partner and the amount invested during the game with a "far" partner. Standard deviation in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01.

While the conflict is complicated by several other actors, the main factions are the Houthi government in Sana'a, which controls the northern part of the country, and the Hadi government, which controls the southern and eastern governorates of former South Yemen. The distinction between "Southern" and "Northern" governorates used here is based on the areas controlled by each faction in 2016. The North-Central & Tihama region was controlled by the Houthi government, but is closer geographically and culturally to the Southern governorates than to the other Northern governorates.

Pairings between Southern (S) and Northern (N2) or Southern (S) and North-Central & Tihama (N1) governorates showed significantly more evidence of near preference than pairings between Northern (N2) and North-Central & Tihama (N1) governorates, the base category in the analysis (Table 4.2.1). All of the regressions control for the player's own region, so we identify the impact of the pairing based on players in each region having been randomly paired with one of the two other regions. As expected, the result is robust to controlling for the order of the games, given that the game order was randomized.

The amount played on average is also lower for these Northern-Southern pairing types (column 3) and most of the change in Near Preference is driven by a decrease in the amount played with the far partner (column 4). This result clarifies that the increased Near Preference is showing overall lower trust in the outgroup, rather than just a bigger difference between the near and far partners.

4.3 Changes in Measured Cooperativeness

We find that Near Preference declined significantly between the baseline and follow-up round of the games, suggesting a positive impact of the program on inter-regional trust. Table 4.3.1 shows that on average Near Preference was lower in follow-up, although the difference is only marginally statistically significant. Column 1 shows both old and new participants, while column 2 shows new participants only. Old participants had higher trust levels at baseline but did not have a significantly lower impact of the program from new participants – if anything, the impact was higher for old participants. This shows that increased exposure to program training, experience, and contacts can continue to generate increased trust even for participants who have already had some exposure to the program.

When we break down the impact of participants by group assignment in column 3, we find a significant decrease in Near Preference among pairs between Northern and Southern governorates,

which had the highest average Near Preferences at baseline. The magnitude of the reduction corresponds to about half of the baseline difference between a S_N2 and N1_N2 pairings.

Table 4.3.1 Changes between baseline and follow-up in main sample

	All participants (1) Near preference	New participants only (2) Near preference	(3) Near preference
Follow-up	-130.8 (87.09)	-128.3 (84.24)	
Southern and Northern pairing (S_N2) * Follow-up			-235.8** (103.15)
Southern and North-central & Tihama pairing (S_N1) * Follow-up			94.3 (122.39)
Northern and North-central & Tihama pairing (N1_N2) * Follow-up			-202.0 (143.61)
Old participant	-339.7** (152.59)		
Follow-up * Old participant	139.9 (239.17)		
Own region fixed-effects?	Yes	Yes	Yes
Observations	564	495	495
Mean dependent variable, Yemeni riyals	49.0	73.8	73.8

Source: Authors' analysis. Note: "Near preference" is defined as the difference between the riyals invested by the respondent during the game with a "near" partner and the amount invested during the game with a "far" partner. Standard deviation in parentheses.
* p < 0.10, ** p < 0.05, *** p < 0.01.

5. CONCLUSIONS

The RUAWFD youth employment program implemented by SFD was designed to both further development goals in rural villages and to provide training for the university graduates selected as youth promoters. Our findings from the survey analysis show that participants perceived RUAWFD training and participation to have increased their future job prospects, with significant positive findings on a range of measures related to their beliefs about their future employability and economic situation.

Secondly, we find that program participation was associated with increasing trust in local institutions as well as increased awareness about marginalized groups. This result is likely to have been aided by the human-rights based approach used in the training of participants.

Surprisingly, we do not find improvements in measured self-efficacy or views on women's role in society. This may be because of the negative time-trends due to the current situation in Yemen. However, it may also be important to try to reinforce the program impact on these outcomes by concentrating more on these topics during training sessions.

Beyond gains to participants themselves, we find evidence that participation in the RUAWFD program was associated with increases in self-reported trust in people outside of the immediate family. This increase in self-reported trust supports our further finding of increases in trust as measured using experimental games. We find a significant decline in Near Preference, our outcome variable which measures the degree to which game players trusted geographically closer partners more than farther partners. This decline is concentrated among partners paired between the

Southern and Northern governorates, i.e. those which are most strongly identified with one of the two main axes in the current civil conflict.

These findings on trust suggest that in addition to increasing the employability of participants, the program can serve as a mechanism for strengthening inter-regional trust levels, an important ingredient for future reconstruction and post-conflict recovery. The results also suggest that it would be useful to make connecting youth participants and former participants from different regions an explicit focus of the program, such as by requiring more active engagement with the WhatsApp group for all program participants, with the objective of building on the foundation of increased trust generated already.

APPENDIX: ATTRITION ANALYSIS

This appendix analyzes the difference in sample size and attrition between the baseline and follow-up surveys, as well as between the baseline survey and participation in either of the two rounds of games.

Attrition in Survey Participation

Appendix Table 1 shows attrition in survey participation. Among the new participants who participated in the baseline survey, 93 percent also submitted a follow-up survey. With the exception of one district, Attayal in Sana'a governorate, where security issues were responsible for a lack of follow-up surveys, attrition was mostly randomly distributed geographically across the sample. The most common reason for survey attrition among new participants was dropping out of the program after filling the baseline survey. Among previous participants in this program, attrition was much higher as they were both harder to reach and the survey team made less of an effort during data collection to collect a follow-up survey from previous participants. Only 68 percent of old participants who answered the baseline survey filled out a follow-up survey. As noted above, only new participants were used for the panel survey analysis presented in this report.

Appendix Table 1 Sample size by survey rounds

	New participants	Previous participants
Male	428 29 missing at follow-up (6.8%) – 21 dropped out of program for idiosyncratic reasons – 3 empty or missing questionnaires – 5 no notes	107 42 missing at follow-up (39.3%) – 12 dropped out of program (worked less than 20 out of 50 days) – 29 not resurveyed because of difficulty relocating – 21 from Attayal (Sana'a) and 8 from Zaibeen (Amran) – 1 no notes
Female	415 26 missing at follow-up (6.3%) – 17 dropped out of program for idiosyncratic reasons – 3 empty or missing questionnaires – 6 no notes	95 23 missing at follow-up (24.2%) – 2 from Wassab Al-Safil (Dhamar) excluded from program due to conflict of interest – 7 dropped out of program (worked less than 20 out of 50 days) – 13 not resurveyed because of difficulty relocating – 11 from Attayal (Sana'a); 2 from Zaibeen (Amran) – 1 no notes
Total surveyed at baseline	843	202
Total resurveyed in follow-up survey	788	137
All observations	1,631	339

Source: Authors' analysis. Note: Idiosyncratic reasons for dropping out of the program include excusing from duty without stating reasons, sickness of the participant, sickness of a close person, marriage or giving birth (for females), absences from the second round and unreachable by phone, travelling in general, travelling for work, getting another job, and conflicts of interest, e.g., having relatives working in the program.

Appendix Table 2 shows the results for testing for differential attrition by baseline characteristics. Overall, we see that few variables were significantly different. As a robustness check, as described in footnote 2, we controlled for characteristics that differed significantly at baseline – trust in local governorate and religious leaders; the perception of cooperation in the respondent's community; the degree to which people are generally concerned with the community welfare; the respondent's ability in planning development projects; and the presence of people in the respondents extended network who were internally displaced because of the current conflict. However, the results did not notably change from those presented in this working paper.

Appendix Table 2 Attrition analysis for the survey panel

	p-value		p-value
Assignment strata	0.24	Likelihood that specialization is required in the market	0.38
Governorate of residence	0.33	Likelihood that your specialization is required at stable time	0.24
Gender	0.76	Want to buy or build your house	0.43
Marital status	0.44	Likelihood of buying or building your house	0.57
Who do you live with?	0.65	Likelihood of buying or building your house at stable time	0.89
Owning a car	0.23	Want to buy/ have car	0.55
Owning a private dwelling	0.98	Likelihood of buying car	0.99
Have done any economic activity during the last 5 years	0.90	Likelihood of buying car at stable times	0.85
Number of activities, mean	0.95	Cooperation in nearby villages	0.72
Employment status (1st work)	0.57	Cooperation in own village	0.05 *
Employment status (2nd work)	0.84	Desired cooperation in own village	0.22
Ever lived outside home village	0.71	Optimism for Yemen in 5 years	0.61
Ever lived abroad	0.96	Optimism for family in 5 years	0.85
Like to live abroad	0.38	Optimism for self in 5 years	0.45
Need permission to spend money	0.32	Conflict effect on self: Felt unsafe	0.29
Have a bank account	0.56	Conflict effect on self: Lost job	0.37
Friends or relatives were community leaders	0.17	Conflict effect on self: Lost property	0.85
Friends or relatives are community leaders	0.19	Conflict effect on self: Left home	0.64
Close friends/ relatives that disagree with politically	0.15	Conflict effect on a close person: Lost property	0.53
Facebook friends that disagree with politically	0.84	Conflict effect on a close person: Left home	0.72
Trust in people in general	0.77	Conflict effect on a close person: Injured or killed	0.29
Trust in your family	0.56	Conflict effect on network: Lost property	0.37
Trust in people in village	0.99	Conflict effect on network: Left from home	0.15
Trust in people in district	0.16	Conflict effect on network: Injured or killed	0.33
Trust in people in governorate	0.37	Know anyone who has been internally displaced	0.06 *
Trust in older generation	0.79	Efficacy Index	0.34
Trust in young generation	0.18	Ability to organize meetings	0.86
Trust in the SFD	0.15	Ability to plan projects	0.03
Trust in tribes	0.13	Ability in managing teamwork	0.63
Trust in local government	0.03 **	Ability in managing conflicts	0.46
Trust in NGOs	0.72	Agree that that people look out mainly for their own interests	0.07 *
Trust in religious leaders	0.02 **	Community contributes to projects	0.40
Trust in news media	0.82	Non-direct beneficiaries in community contribute to projects	0.52
Trust in senior government officials	0.63	Active cooperative organizations in community?	0.92
Trust in leaders	0.15	Community development projects that could be implemented in community?	0.37
Trust in political parties	0.75	Marginal groups not included in community decisions	0.88
Want to have permanent job	0.35	Young people included in community decisions	0.38
Likelihood of permanent job	0.43	Female Empowerment Index	0.48

Source: Authors' analysis. Comparison of baseline characteristics for resurveyed vs. non-resurveyed among new participants. Note: p-values in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01.

Attrition in Game Participation

Appendix Table 3 summarizes the size of the original sampling frame and the actual sample sizes used for analysis of the games. In total, 527 participants were invited to play in the first round of the games and 599 were invited in the second round. However, the sample used for analysis is smaller due to non-response and concerns about maintaining balance.

Appendix Table 3 Sample size for games

	Baseline				Follow-up				Total
	Total	South- ern	North- ern	North- Central & Tihama	Total	South- ern	North- ern	North- Central & Tihama	
Original sampling frame for games (follow-up sampling frame includes additions from baseline survey)	527	125	234	168	599	225	222	152	1,126
Game participants that have baseline survey data	403	105	173	125	599	225	222	152	1,002
<i>Have baseline survey and assigned to group correctly</i>	292	100	117	75	415	170	141	104	707
<i>Have baseline survey and responded to both games</i>	341	90	139	112	447	156	178	113	788
Assigned to group correctly and responded to both games	258	87	103	68	324	124	118	82	582

Source: Authors' analysis.

The baseline games were conducted before the baseline survey. Consequently, the exclusion of participants who dropped out of the program before responding to the baseline survey reduces the baseline games sample size from 527 to 403. Additionally, due to uneven group sizes, some participants were unable to be properly matched with both a “near” and “far” partner. A further reduction in sample size was caused by a mistake in coding some of the governorate locations where the text field was left blank. Excluding participants who were not properly matched due to these reasons results in a sample size of 292 in baseline games and 415 in follow-up games. Finally, not all participants responded to the text message game in spite of being contacted individually by phone over a period of several days in cases of non-response. In the baseline games 7 percent of participants never responded, while in the follow-up games 15 percent of participants never responded. Some participants responded only for one of the two games that they were asked to play. The total share that did not respond to both games in the first round of games was 15 percent and in the second round was 25 percent.

Our final sample size using the three randomly assigned groups is 258 participants at baseline and 324 at follow-up. This is the sample size used when using the outcome variable “Near Preference” and showing the difference in outcome by group assignment.

Our checks for attrition bias were as follows. By restricting the sample to only participants for whom we have baseline survey data, we are able to assure that baseline characteristics are still balanced in spite of this high rate of attrition.

Because our final analysis sample is considerably smaller than the planned sample due to high non-response rates and some issues with the randomization, we are concerned, first, about the balance between the baseline and follow-up groups and, second, about the balance conditional on own location by group assignment. However, Appendix Table 4 shows that across the large number of baseline variables checked, there are few significant differences between the participants who played in the first and second round of the games. One significant difference is that self-reported trust levels were higher among baseline players. Baseline players were more trusting of family,

village, and older generation, more likely to want permanent job, perceived higher cooperation in own village, were more optimistic about the situation of Yemen after five years, and perceive more cooperation in nearby communities. This shows that the samples are overall not too poorly balanced and that, if anything, our results, which show higher trust levels in follow-up games, are biased downwards by the imperfect balance.

We also repeated the same balance tests for the partner randomization conditional on own location. For the balance between participants in the Southern governorates paired with Northern or North-Central & Tihama governorates, we find 7 out of 90 baseline variables significant at the 10 percent level or less. For the balance between participants in North-Central & Tihama governorates paired with Southern versus Northern governorates, we find 12 out of 90 baseline variables significant at the 10 percent level or less. For the balance between participants in Northern governorates paired with Southern versus North-Central & Tihama, we find 6 out of 90 baseline variables significant at the 10 percent level. These results are in line with the number of variables that would randomly be expected to show significant differences at the 10 percent level.

Appendix Table 4 Balance test for participants in first vs. second round of games

	p-value		p-value
Gender	0.94	Likelihood of permanent job	0.35
Days worked	0.17	Likelihood that specialization is required in the market	0.35
Amount paid	0.53	Likelihood that your specialization is required at stable time	0.86
Highest qualification	0.32	Want to buy or build your house	0.58
Marital status	0.44	Likelihood of buying or building your house	0.18
Who do you live with?	0.07	Likelihood of buying or building your house at stable time	0.59
Owning a car	0.26	Want to buy/ have car	0.16
Owning a private dwelling	0.56	Likelihood of buying car	0.42
Have done any economic activity during the last 5 years	0.74	Likelihood of buying car at stable times	0.72
Number of activities, mean	0.22	Cooperation in nearby villages	0.09 *
Employment status (1st work)	0.09 *	Cooperation in own village	0.13
Employment status (2nd work)	0.31	Desired cooperation in own village	0.22
Ever lived outside home village	0.55	Optimism for Yemen in 5 years	0.01 **
Ever lived abroad	0.96	Optimism for family in 5 years	0.26
Like to live abroad	0.52	Optimism for self in 5 years	0.27
Need permission to spend money	0.64	Conflict effect on self: Felt unsafe	0.40
Have a bank account	0.18	Conflict effect on self: Lost job	0.77
Friends/ relatives were community leaders	0.27	Conflict effect on self: Lost property	0.65
Friends/ relatives are community leaders	0.99	Conflict effect on self: Left home	0.16
Close friends or relatives that disagree with politically	0.62	Conflict effect on a close person: Lost property	0.29
Facebook friends that disagree with politically	0.06	Conflict effect on a close person: Left home	0.12
Trust in people in general	0.20	Conflict effect on a close person: Injured or killed	0.66
Trust in your family	0.05 *	Conflict effect on network: Lost property	0.16
Trust in people in village	0.05 *	Conflict effect on network: Left from home	0.64
Trust in people in district	0.60	Conflict effect on network: Injured or killed	0.30
Trust in people in governorate	0.16	Know anyone who has been internally displaced	0.14
Trust in other areas of Yemen	0.78	Efficacy Index	0.41
Trust in older generation	0.02 **	Ability to organize meetings	0.27
Trust in young generation	0.86	Ability to plan projects	0.53
Trust in the SFD	0.62	Ability in managing teamwork	0.80
Trust in tribes	0.29	Ability in managing conflicts	0.39
Trust in local government	0.24	Agree that that people look out mainly for their own interests	0.20
Trust in NGOs	0.28	Community contributes to projects	0.22
Trust in religious leaders	0.99	Active cooperative organizations in community?	0.91
Trust in news media	0.21	Community development projects that could be implemented in community?	0.41
Trust in senior government officials	0.92	Marginal groups not included in community decisions	0.18
Trust in leaders	0.42	Young people included in community decisions	0.66
Trust in political parties	0.29	Female Empowerment Index	0.79
Want to have permanent job	0.07 *		

Source: Authors' analysis. Note: p-values in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01.

REFERENCES

- Bauer, M., A. Cassar, J. Chytilová, and J. Henrich. 2014. "War's enduring effects on the development of egalitarian motivations and in-group biases." *Psychological Science* 25 (1): 47-57.
- Colletta, N.J., and M.L. Cullen. 2000. *The nexus between violent conflict, social capital and social cohesion : Case studies from Cambodia and Rwanda*. Social Capital Initiative Working Paper No. 23. World Bank, Social Development Family, Environmentally and Socially Sustainable Development Network. Washington, DC: World Bank.
- Collier, P., V.L. Elliott, H. Hegre, A. Hoeffler, M. Reynal-Querol, and N. Sambanis. 2004. *Breaking the Conflict Trap: Civil War and Development Policy*. Washington, DC: World Bank.
- Fearon, J.D., M. Humphreys, and J.M. Weinstein. 2009. "Can development aid contribute to social cohesion after civil war? Evidence from a field experiment in post-conflict Liberia." *American Economic Review* 99 (2): 287-91.
- Gilligan, M.J., B.J. Pasquale, and C. Samii. 2014. "Civil war and social cohesion: Lab-in-the-field evidence from Nepal." *American Journal of Political Science* 58 (3): 604-619.
- Johnson, N.D., and A.A. Mislin. 2011. "Trust games: A meta-analysis." *Journal of Economic Psychology* 32 (5): 865-889.
- Pettigrew, T.F., and L.R. Tropp. 2006. "A meta-analytic test of intergroup contact theory." *Journal of Personality and Social Psychology* 90 (5): 751.
- Woolcock, M., and D. Narayan. 2000. "Social capital: Implications for development theory, research, and policy." *The World Bank Research Observer* 15 (2): 225-249.

ABOUT THE AUTHORS

Olivia Bertelli is an Assistant Professor in LEDa (Laboratory of Economics Dauphine) at Université Paris-Dauphine, based in Paris, France. **Sikandra Kurdi** is an Associate Research Fellow in the Egypt Strategy Support Program (ESSP) of the International Food Policy Research Institute (IFPRI), based in Cairo. **Mai Mahmoud** at the time of writing was a Senior Research Assistant in ESSP of IFPRI, based in Cairo and is now a PhD student at Tufts University, Medford, Massachusetts, USA. **Mohamed Al-Maweri** and **Tareq Al Bass** are monitoring and evaluation officers for the Social Fund for Development, based in Sana'a, Yemen.

ACKNOWLEDGMENTS

The impact evaluation project was managed by the World Bank as part of the Yemen Emergency Crisis Response Project and was supported by the Nordic Trust Fund. Lamis Al-Iryani, the head of Monitoring and Evaluation for the Social Fund for Development, managed the overall coordination of this project.

From the World Bank, we gratefully acknowledge the guidance and support provided by Afrah Al-Ahmadi, continuing support for the evaluation from Qaiser Khan, and valuable comments and suggestions from Yashodhan Ghorpade.

Gamdan Al-Ebssy of the Information Technology Unit of the Social Fund for Development coded the TeleRivet application to send SMS messages for playing the games and recording the outcomes. We are also thankful for the contributions of several experts from the Social Fund for Development for assisting with the data collection and data entry processes, notably Abdul Karim A. Waheish, Eisa T. Alhetar, and Abdullah H. Al Hassani.

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

1201 Eye St, NW | Washington, DC 20005 USA
T. +1-202-862-5600 | F. +1-202-862-5606
Email: ifpri@cgiar.org | www.ifpri.org | www.ifpri.info

IFPRI-EGYPT

World Trade Center, 1191 Corniche El Nile, Cairo, Egypt
T: +20(0)225778612
<http://egyptssp.ifpri.info/>

The Middle East and North Africa Regional Program is managed by the Egypt Strategy Support Program (Egypt SSP) of the International Food Policy Research Institute (IFPRI). The research presented here was conducted as part of the CGIAR Research Program on Policies, Institutions, and Markets (PIM), which is led by IFPRI. This publication has been prepared as an output of Egypt SSP. It has not been independently peer reviewed. Any opinions expressed here belong to the author(s) and do not necessarily reflect those of IFPRI, PIM, or CGIAR.

© 2019, Copyright remains with the author(s). This publication is licensed for use under a Creative Commons Attribution 4.0 International License (CC BY 4.0). To view this license, visit <https://creativecommons.org/licenses/by/4.0>.