Stay in the Game:

A Randomized Controlled Trial of a Sports and Life Skills Program for Vulnerable Youth in Liberia

LORI BEAMAN

Northwestern University

SYLVAN HERSKOWITZ

International Food Policy Research Institute

NIALL KELEHER

Innovations for Poverty Action

JEREMY MAGRUDER

University of California, Berkeley

February 2020

Abstract

Over the past two decades, sports programs have proliferated as a mode of engaging youth in development projects. Thousands of organizations, millions of participants, and hundreds of millions of dollars are invested in sports-based development programs each year. The underlying belief that sports promote socioemotional skills, improve psychological well-being, and foster traits that boost labor force productivity has provided motivation to expand funding and offerings of sport for development (SFD) programs. We partnered with an international NGO to randomly assign 1200 young adults to a sports and life skills development program. While we do not see evidence of improved psychosocial outcomes or resilience, we do find evidence that the program caused a 0.12 standard deviation increase in labor force participation. Secondary analysis suggests that the effects are strongest among those likely to be most disadvantaged in the labor market.

1   Introduction

Across the world, youth are found playing sports wherever there is enough space to kick a ball or swing a bat: on fields, beaches, empty streets, or vacant lots. Few activities capture the excitement, energy, and imagination of youth as much as sports. Beginning in the 2000s, a wave of exuberance brought increased popularity to the use of sports for the promotion of development outcomes. "Sport for Development" (SFD) programs take many forms but broadly aim to develop pro-social behaviors and labor force potential among young adults, with a frequent emphasis on at-risk and difficult to reach populations.

In 2001, the United Nations established the Office on Sport for Development and Peace. Four years later they declared 2005 the International Year of Sport and Physical Education and defined Sport for Development as a broad effort to "engage people from disadvantaged communities in physical activity projects that have an overarching aim of achieving various social, cultural, physical, economic or health-related outcomes."[1] A recent review of SFD programs found 955 organizations engaging exclusively in sports programming and over two thousand additional organizations that incorporate sports into their programs (Svensson and Woods 2017). Other sources corroborate these estimates, claiming that SFD programs reach tens of millions of youth throughout the world (Adair 2014). Our own efforts to track down expenditures on SFD programs suggest that

[1]We follow this definition from Adair (2014) to preserve a focus on sports programming as a form of intervention. Sports for Development and Peace, Sport for Change, Development through Sport, and Sport and Development are other commonly used terms for the types of programs we discuss in this paper.

global expenditures exceed hundreds of millions of dollars per year.[2] Both in terms of participation and expenditure, SFD programs operate at a massive global scale.[3]

SFD proponents frequently tout sports-centered programs as an impactful form of direct intervention for at-risk youth as well as an effective entry point for complementary programs targeting difficult to reach populations. In particular, it is frequently asserted that these programs can improve psychosocial outcomes and soft skills of participants and that these benefits will, in turn, lead to better labor market outcomes for marginalized youth.[4] Despite this large mobilization of resources, there is little existing evidence to either substantiate or refute these claims. Unlike interventions involving labor force training, there is little experimental evidence to guide policymaking in the SFD sector.

In Liberia, the site of this study, over three-fifths of the country's population is below the age of twenty-five. This foreshadows continued rapid population growth and presents considerable demographic challenges engaging such a large youth population going forward.[5] Although the country's second civil war ended in 2003, the memories and trauma of one of the continent's bloodiest and most protracted conflicts still looms large on the national psyche. Policy-makers and international actors have therefore been worried that failure to engage with youth and help them find productive outlets would risk destabilizing the country and eruption of civil unrest.

---

[2] This estimate is based on reports and documentation found online. We examined organizations that self-register through `https://www.sportanddev.org` and `http://www.streetfootballworld.org`. The eight members of the steering committee of the International Platform for Sport and Development alone could be linked to over $105 million in project budgets for SFD-related activities. Searching for additional financial disclosures led us to $219.5 million in annual expenditures. This is likely a considerable under-counting of overall SFD expenditures. Further details are provided in Appendix B.

[3] If SFC expenditures are between $300–600m per year, this would be comparable to 0.3-0.55% of all global aid given by the OECD's 30 Development Assistance Committee (DAC) members and equivalent or greater than nine of these countries' total contributions (Organization for Economic Cooperation and Development 2019). According to a recent World Bank publication, it is also close to the total amount spent in 2014 on non-communicable diseases (Bendavid et al. 2017).

[4] Many organizations claim that these benefits are even wider ranging with the United Nations asserting that benefits include gender equality, social integration, development of social capital, peacebuilding, conflict prevention and resolution, trauma relief, and economic development, `https://www.un.org/sport/content/why-sport/overview`

[5] Liberia's 2014-2015 Household Income and Expenditure Survey shows that 76.4 percent of Liberia's population is under the age of 35, 62.5 percent is under 25, and 44.5 percent is under 15.

In this paper, we investigate the effect of a life-skills program mediated through sports groups, on psychosocial well-being and labor force outcomes among youth in Monrovia, Liberia's largest city. Our evidence derives from a randomized control trial of an SFD program conducted by Mercy Corps, a large international non-governmental organization with similar programs in over twenty-five countries. The study took place in nine communities around Monrovia, with 1,200 youth invited to participate in the Mercy Corps program and another 1,200 assigned to the control group. Similar to the claims of many other SFD proponents and practitioners, the stated aims of Mercy Corps' SFD program, "Sports for Change" (SFC), are to improve psychosocial or socio-emotional behaviors among participants, increasing their resiliency and "readiness" for productive labor-force participation. The SFC method centers on the formation of youth groups, engagement with these groups through competitive sports, and facilitation of a complementary life-skills curriculum. Random assignment of individuals to youth groups and a control group allows us to estimate the causal impact of SFC on a range of psychosocial and labor market outcomes.

Anticipating benefits on psychosocial outcomes, our data collection was designed to allow construction of five measures of psychosocial outcomes, which we aggregate into an overall psychosocial index (PSI). Corroborating the implementer's theory of change, we see a strong positive baseline correlation between psychosocial well-being and labor force measures. However, our analysis reveals limited evidence of meaningful direct impacts of the program on psychosocial outcomes, with a point estimate of -0.014 standard deviations and a 95% confidence interval ranging from -0.098 to 0.07. We also do not find any evidence of improved resiliency among program beneficiaries faced with challenging life events.

The proposed theory of change suggested that psychosocial improvements caused by the program would lead to improved labor force outcomes. And yet, despite lack of measurable psychosocial impacts, we do see a statistically significant increase in labor

outcomes. Our results show an increase in our aggregate labor force index among study participants with a point estimate of 0.12 standard deviations and a 95% confidence interval from 0.03 to 0.20 standard deviations. Point estimates for labor supply and earnings both suggest increases of approximately 12%.

Given the lack of evidence of impacts on psychosocial measures and as a hypothesized causal pathway, we examine whether peer effects may have contributed to results on labor market and psychosocial outcomes. We find limited evidence to support the importance of peer effects in this context, unable to explain the results through exposure to high (or low) performing peers or the presence of pre-existing social ties in assigned groups. Finally, we explore heterogeneity of the treatment effects in order to better understand who is benefiting from the program. It appears that along all the dimensions that we tried (young, uneducated, female, un-trained), labor force benefits of the program are bigger for each of these marginalized populations.

2   Related Work

Sport for Development proponents assert a wide range of benefits for program participants, arguing that sports are an effective way to improve pro-social and psychosocial "life skills" as well as labor outcomes for participating youth. Many organizations trust that benefits result directly and naturally from sports participation: sports may inherently teach valuable life lessons and foster pro-social skills. As a result, some organizations focus exclusively on expanding opportunities for youth to play sports, such as through broad distribution of soccer balls in developing countries. Other organizations have invested considerable time and effort designing and incorporating complementary life skills training into their sports activities.[6] The body of existing evidence on SFD pro-

---

[6]See Ball to All (https://www.balltoall.org/) as an example of a sports only program and Peace Players International (https://www.peaceplayers.org/) and Grassroots Soccer (https://www.grassrootsoccer.org/) as examples of programs with complementary soft skills programming.

grams of any form is limited. A large portion of the existing literature focuses on improving the design and strategic implementation of SFD programs (e.g Hartmann and Kwauk 2011; Jeanes 2013).

Alongside these prescriptive suggestions for SFD are implicit claims about expected positive impacts on participants. In one example, Kidd (2008) states "[SFD] has brought considerable benefit to many children and youth in the countries where it is conducted." Rookwood (2008) suggests that soccer builds trust, respect, and self-discipline. Others highlight the positive relationship between sports and psychosocial development and resiliency (Petitpas et al. 2005; Berlin et al. 2007; Perkins and Noam 2007; Henley et al. 2007). Frequently studied examples of the SFD movement range from Nairobi to the United States, broadly framing SFD programs as a mechanism to prepare young adults to utilize pro-social skills in a variety of long-term applications.[7]

While these studies provide important psychological and sociological grounding of hypothesized effects and potential benefits, the evidence relies heavily on case studies and theoretical models unsubstantiated by empirical evidence. Three recent reviews of this literature have agreed that the evidence base is lacking in rigor and that the few studies with individual-level data are systematically under-powered in sample size and fail to address concerns of endogeneity between program participation and program outcomes (Holt and Jones 2007; Burnett 2009; Coalter 2010).

Since these reviews, there have been a handful of recent studies using more rigorous empirical methods. Panter-Brick et al. (2018) assess the impact of another Mercy Corps program in Syria similar their Sports for Change program in Liberia. Jordanian and Syrian youth were organized into groups centered on a range of activities (often, but not exclusively sports based) while locally identified facilitators provided life skills and psychosocial support. The authors found modest improvements in some psychosocial

---

[7]For example, the Mathare Youth Sports Association in Nairobi (Morris et al. 2004; Kidd 2008; Coalter 2009), the Physically Active Youth program in Namibia (Donnelly et al. 2007), and the Midnight Basketball Leagues across numerous locations in the United States (Hartmann and Depro 2006)

outcomes with bigger effects among youth exposed to higher levels of trauma. However, the study was hindered by high levels of attrition (74% after one year) and did not examine economic outcomes. Three other recent studies also focused on sports programming as a way to promote reconciliation across groups. In India, Israel, and Iraq, these studies found that participating in sports increased pro-social behavior resulting from cross-caste, ethnic, or religious interactions (Ditlmann and Samii 2016; Mousa 2018; Lowe 2018).

This paper also relates to the broader body of work linking non-cognitive and psychosocial factors to labor market outcomes. Two recent surveys of this literature document how cognitive skills do not fully explain labor market outcomes, attributing a large part of the unexplained variation to socio-emotional skills (Heckman and Kautz 2012; Kautz et al. 2014). These "character," or life skills, are seen as necessary for the realization of investments in cognitive skills and the authors suggest that interventions that improve these soft skills may have an impact on economic and welfare outcomes as well.

In response to this body of work linking soft skills and economic outcomes, considerable interest and programming have been put into psychosocial interventions in traumatized populations, although findings have been inconsistent (Underwood 2018). On the positive end, Adoho et al. (2014) present evidence from a randomized control trial of an intensive program on economic empowerment and life skills training for young Liberian women. They find large improvements in psychosocial measures of self-confidence and anxiety, a 47% increase in employment, and an 80% increase in earnings for program participants. Ibarraran et al. (2014) also have positive, but more modest, results on the impacts of a soft skills training program for youth in the Dominican Republic, finding improvements in earnings (8%), formality of work (10%), and measures of non-cognitive skills (0.08-0.12 standard deviations), but no increase in overall labor force participation. Calero and Rozo (2016) have mixed results in another experimental study among youth in Brazil's Favelas, finding that training aimed at reducing risky behaviors increased income but only improved behaviors among participants with higher pre-existing levels of

socio-emotional skills. Groh et al. (2016) find no effects of a soft skills training program on female youth employment among Jordanian community college graduates. And in the most similar setting to our paper, Blattman, Jamison, and Sheridan (2017) provide mixed evidence in their study of pro-social programming in a sample of high-risk ex-combatants in Liberia. The authors find that a cognitive behavioral training (CBT) intervention led to a 0.25-0.3 standard deviation decrease in antisocial behaviors one month after completing the CBT program. However, this effect only persisted in follow-ups one year after the intervention among CBT beneficiaries who also received a substantial cash grant of USD $200. They find no persistent impacts of the program on economic outcomes.

This paper bridges the literature on sports and psychosocial behavior with the literature on psychosocial training as related to labor market outcomes. In the context of our study, we examine a setting different from that examined in other sports program evaluations. Namely, the youth population in Monrovia is less segregated than that in Lowe (2018) and not immediately effected by national conflict as in Ditlmann and Samii (2016) and Mousa (2018). We build on this literature by looking at longer-term impacts of an SFC program on a broader set of psychosocial measures (approximately one year after the intervention). We also leverage a larger sample of study participants randomized into sports groups. In addition, the data we collected allows us to examine labor force outcomes as well as psychosocial outcomes. While theory driving the SFD literature presupposes the connection between sports and psychosocial outcomes, it is not a given, nor is it obvious that changes in labor force outcomes are predicated on changes in psychosocial outcomes. In this paper, we present evidence intended to unpack the linkages between sports programs, psychosocial outcomes, and labor force outcomes.

3   Context and Program Design

At the time of the study, the President of Liberia, Ellen Johnson-Sirleaf, expressed concern over high levels of youth unemployment as a destabilizing factor for the country (Dunmore 2013). Overall labor force participation was and remains very low in Liberia, estimated by the International Labor Organization at roughly 60%.[8] The central role of youth as both victims and combatants in the Liberian conflict amplifies a sense of urgency among policy-makers and international actors to find ways to engage young populations in positive activities as well as to help those directly and indirectly affected by the war (see Appendix Section F for a summary of the qualitative research).

With this backdrop, Mercy Corps launched the Promoting Sustainable Partnerships for Economic Transformation (PROSPECTS) initiative in 2012 in Montserrado County of Liberia, where Monrovia is located. In this paper, we evaluate the impact of one program within PROSPECTS, namely the Sport for Change program.

Sport for Change (SFC) targeted vulnerable, out-of-school youth between the ages of 15 and 25 with little or no prior formal work experience. This target population was broadly regarded as unskilled and "unemployable." The SFC program in Liberia was designed to use sports groups as a means of attracting and engaging vulnerable youth to participate in pro-social activities. The SFC methodology, as designed by Mercy Corps, seeks to leverage the beneficial potential of team-based sports with life skills training. As such, Mercy Corps integrated five core life skills into the SFC program: (1) resilience, (2) strategy making and planning, (3) teamwork and trust building, (4) self-esteem, and (5) constructive communication (see Appendix Section D for details of the SFC program and Appendix Figure A.1 for an example of a session schedule). The methodology closely mirrored the international SFC approach Mercy Corps has implemented in more than 25 countries to engage youth in post-conflict settings.[9] Bundling sports team practices with

---

[8] http://data.un.org/en/iso/lr.html

[9] For a description of the Sport for Change approach deployed internationally by Mercy Corps, see https://www.mercycorps.org/tags/sport-for-change.

life-skill activities, Mercy Corps expected positive impacts on participants' psychosocial outcomes and improved resiliency to adverse life events. Improved psychosocial well-being and resiliency were expected to create a foundation of workforce "readiness" for participants to enter into formal employment or launch their own small businesses.

Contemporaneous to the timing of this study, Liberia experienced a national emergency as a result of an outbreak of Ebola in West Africa. Although the first cases of Ebola were reported in Liberia in March 2014, the SFC program had concluded before the first recorded Ebola deaths were documented in Monrovia. Over the course of the following year, a total of 10,675 cases of Ebola in Liberia with 4,809 resulting in deaths.[10] Nearly one-third of these deaths occurred in Montserrado County, where this study was implemented. While the Ebola outbreak did not affect implementation of the program, in the months following the end of SFC, it had a considerable impact on daily life in Monrovia and the crisis was ongoing at the time of the endline.

4   Experimental Design and Data

Recruitment and Random Assignment

We worked with Mercy Corps in nine urban communities of Monrovia to identify a pool of eligible youth for participation in the Sport for Change program. Public announcements were made in each community, inviting young adults between the ages of 15 and 25 to attend a meeting organized by Mercy Corps to register for potential involvement in the Mercy Corps program and in the study. We organized one recruitment event per community. At the event, applicants were randomly selected to participate in the Mercy Corps program via a public lottery. In order to prevent gender imbalance, we stratified the lottery by applicant gender. Women and men formed separate lines and each indi-

---

[10]Information on Ebola cases from the Center for Disease Control, `http://www.cdc.gov/vhf/ebola/outbreaks/2014-west-africa/case-counts.html`.

vidual drew a ticket indicating their group assignment from a covered bucket and were unable to change their assignment after their group assignment was revealed.

In total, 1,200 individuals were assigned to an SFC team while 1,200 were not invited for the Mercy Corps program and thus would serve as our control group in the study. With labor force outcomes as the primary outcome of interest, power calculations showed that we our design allows us to detect a small minimum detectable effect of 0.16 standard deviations with 80% power, assuming that the standard deviation of treatment effects across communities is 0.1 and that baseline covariates explain 20% of the variation in final outcomes.

At the enrollment events, registration forms were completed by the entire pool of eligible applicants regardless of group assignment at the registration events in each community. Registration staff recorded basic demographic information on age, gender, and schooling, plus extensive tracking and contact information in anticipation of the baseline interviews. Additional details on the recruitment of participants into the study and random assignment to youth sports groups can be found in Appendix C.[11]

Registration events proceeded sequentially one community at a time. The initial event occurred in the West Point community in July 2013 and the final event was held in the Logan Town community in February 2014. Table 1 shows the timing of the registration event, baseline survey, and the number of study participants by community. The number of registered participants varied by community from 160 to 480 depending on the anticipated size of its youth population. Mercy Corps sought 50% female participants.

---

[11]The initial research and program design of PROSPECTS also included a "cash for work" (CFW) program that involved an opportunity for participants to earn money by collecting recyclable materials. Implementation challenges, compounded by low interest in the program, meant that very few people assigned to the CFW group ever actually participated in the program. In total, 1,200 individuals were assigned to the cash for work program. Of those assigned to CFW, 976 attended at least one cash for work meeting and only 467 attended more than two meetings. As a result, the analysis in this paper focuses solely on the Sport for Change component of PROSPECTS. Random assignment of the full set of 3,000 participants in the original research design assigned 20% of participants to the CFW program only, 20% to the SFC program only, 20% to both programs, and 40% to the control group. Due to implementation challenges as well as budgetary and logistical constraints on conducting follow-up surveys amid the outbreak of Ebola, it was decided to exclude respondents in the CFW-only treatment from the endline survey.

Although there was some variance of female participation levels across communities, gender balance across treatments within each community was broadly preserved through our stratification procedures. Ultimately, women comprised 52.2% of the treatment group and 51.8% of the control group.

Baseline Summary Statistics

In the days immediately following the registration event in each community, Innovations for Poverty Action (IPA) conducted in-person baseline interviews with all registrants following the schedule in Table 1.[12] Between completion of the registration events and administration of the baseline survey, there was very low attrition: only five out of 2,400 registrants refused to be interviewed or could not be found after the registration event.

Table 2 shows summary statistics and balance of participants at baseline by program treatment status. The average age in the sample is 21 years old. 83% had completed primary school and slightly more than 25% had completed secondary. Just over 43% had some form of employment at baseline. Among those working, respondents worked roughly 28 hours per week and earned approximately USD $16.50.

While most baseline variables are well-balanced across treatment status, we note an imbalance in five of the 35 variables which is slightly higher than we would expect from chance. Respondents in the control group had higher baseline measures of self-esteem and numeracy, but lower scores on the depression, anxiety, and stress (DASS) index and three-month income. For these variables, we note that even though differences are statistically significant, the magnitudes of these differences are all less very small and unlikely to be economically meaningful.

The sole exception is for three-month income which has a relatively large difference

---

[12]For logistical reasons, it was not possible to both generate a sample frame of individuals and also complete baseline surveys prior to treatment assignment. This design feature did not lead to differential baseline survey response rates. Of the five registrants who did not participate in the baseline, two were in the treatment group and three were in the control, attrition rates of 0.17% and 0.25% respectively.

in means. However, this baseline imbalance is no longer significant after implementing the inverse hyperbolic sine transformation, which mitigates the influence of outliers in the data with unusually large reported earnings. Still, given that longer recall windows are prone to greater amounts of measurement error, we exclude the three-month income measure from our primary analysis of labor outcomes and give preference to reported seven-day income and show robustness of our main results to including it in the appendix.

Overall, we may have been concerned about whether baseline interviews occurring after assignment to treatment (but before the intervention began) could have affected reported baseline responses. However, the limited extent of baseline imbalance is reassuring and suggestive that, if present, this bias is likely to be small.

Negative Life Events, Psychosocial and Labor Force Measures

The Sport for Change program aimed to improve participants' psychosocial well-being and resiliency. In addition to their direct benefits, consistent with claims in the SFD community and literature, it was believed that improving psychosocial well-being would impact employment and workforce readiness, leading to higher labor force participation and earnings. To assess this motivation, we collected data on negative life events (to explore resiliency), psychosocial measures, and labor force measures. Without distinct predictions for different individual components of these three groups, we construct aggregate indices of each of these measures for each individual.

First, we create an index of negative life events. The surveys asked if respondents or their families had been affected by a set of different types of negative life events over the past year. For example, 27 percent of the control group reported a serious accident that injured a member of the household; 28 percent reported experiences of abuse or a violent crime. Using these responses, we created a life event index by adding together the number of affirmative responses given by the respondent and standardizing this re-

sulting sum. Respondents in the control group and their families had been impacted by an average of 1.5 of these negative life events, with a standard deviation of two.[13] The standardized index is coded so that higher values indicate that a respondent's household experienced more negative events.

Second, we collected data on a large number of questions linked to different psychosocial measures and indices used elsewhere in the literature: subjective welfare, self-esteem, locus of control, aggression, and risky behaviors. Subjective welfare is measured by asking respondents where they see themselves on a six rung ladder with six as the highest possible response. The mean response in the sample was 2.3. Locus of Control is a measure where higher values indicate that a respondent feels that they are more in control of their life outcomes. The self-esteem index measures whether people articulate relatively good or bad feelings about themselves. The aggressive behaviors index captures reported interactions considered to be aggressive, such as disputes with a neighbor or peer. And finally, we collect data in order to form a risky behaviors index including reported gambling, smoking, alcohol, and drug use. All indices are standardized and coded so that positive values reflect "better" behaviors. Additional information on the sub-questions and construction of these indices are included in Appendix D. From these different standardized measures of psychosocial well-being, we then create an aggregate, Psychosocial Index (PSI) by standardizing the sum of these five components.

Labor outcomes are captured in the survey in the form of reported hours worked and earnings. We create a standardized labor force index by standardizing reported hours worked and the inverse hyperbolic sine transformation of earnings (to address skewness in earnings). This sum is then re-standardized and used as the labor force index (LFI) in the main analysis.

The program was motivated by a belief that psychosocial well-being and labor force

---

[13]Some life events recorded in the interview were ultimately excluded from the index where we thought that the likelihood of this event could be affected by the program. Appendix Table A.9 shows the separate events and indicates with a star which measures were included in the life event index.

engagement are closely linked. Table 3 shows a significant and positive relationship between three of the five measures of psychosocial behaviors and the labor force index. The correlation between LFI and the aggregate psychosocial index is particularly strong and highly significant. Jointly estimated associations between the psychosocial measures and LFI are included in Appendix Table A.2 and lead to similar conclusions.

These correlations add credibility to claims of a relationship between psychosocial well-being and labor outcomes, but do not constitute a test of causality between the two (in either direction). Importantly, the correlations suggest that we are capturing meaningful psychosocial measures in our survey which relate to labor market outcomes.

Program Implementation

Mercy Corps began implementation of the SFC program, following completion of a community's baseline. In total, Mercy Corps established 30 unique sports clubs with 40 members per club. For each club, Mercy Corps recruited two coaches from adults living within the community. All coaches participated in a mandatory five-day training organized by Mercy Corps. Training curriculum covered facilitation skills, the SFC methodology, basic first aid, and the responsibilities of coaches. In addition, Mercy Corps hired four coach mentors to provide ongoing support to the coaches throughout the program with continued training, help lesson planning, and assistance problem solving any challenges they were experiencing with their teams. Additionally, Mercy Corps conducted audits of sports club meetings to ensure consistency across SFC groups with Mercy Corps' international standards.

Coaches organized a total of 16 sessions, typically one or two per week. The planned three-hour meeting comprised one hour of introduction and warm-ups, one hour of instructional activities, and one hour of sports, typically soccer and handball. Table 4 presents the topics covered in each of the sixteen Sports for Change sessions along with the targeted skills emphasized in each session. Participants received USD $2 for each

session that they attended, a sum intended to reimburse participants for transportation expenditures.

Participation in SFC was high in all nine communities; 73% of youth assigned to a SFC group attended at least one session. Figure 1 shows that 65% of all SFC-assigned subjects attended at least 80% of their group's meetings. On average, the 1,200 individuals assigned to SFC attended 10.35 SFC sessions (out of a maximum of 16). These high attendance rates suggest that the SFC program was desirable in the eyes of participants and adds credibility to the view that this program was well-implemented. This is corroborated by endline survey responses where 96% of those that attended at least one SFC session said that they enjoyed the program. Respondents also saw the value of the SFC program applying to a wide range of settings. In response to the question, "In what contexts do you think the SFC skills are most useful?" the plurality of respondents, 36%, said when playing sports. However, 21.2% responded that the program was most useful for conflict resolution while 31.5% said that the program was most useful for finding employment. Additional qualitative interviews revealed that participants saw value in the social and life skills training portions of the SFC program. Further details on these are included in Appendix F.

Endline Survey

Following completion of implementation of the SFC program in all communities, we attempted follow-up interviews of all baseline respondents assigned to either the control group or one of the SFC teams. The endline survey was intended to be conducted in person with respondents one year after completion of the intervention. However, due to risks associated with travel restrictions and quarantines during the Ebola crisis in Liberia, the endline survey was administered through computer-assisted telephone interviews.[14] The

---

[14]Garlick, Orkin, and Quinn (Forthcoming) show in a study with microenterprise owners in South Africa that phone based interviews did not reduce data quality of labor outcomes.

endline survey was conducted simultaneously for participants in all nine communities.[15] Endline interviews began on April 3, 2015, and were completed on May 9, 2015.

Despite the challenges of conducting a phone-based survey during the Ebola outbreak, 2,081 individuals were successfully interviewed for the endline survey, a follow-up success rate of 87%. We test for selective attrition in Appendix Table A.1 and find no evidence of selective attrition in terms of total number of attriters by treatment status (columns (1)-(2)). In column (3) we see no overall evidence of selective attrition and the joint test of significance for baseline covariates interacted with treatment has a p-value of 0.98. There is one strongly significant baseline characteristic, baseline psychosocial index (PSI). Participants in the treatment group with low baseline PSI measures were more likely to attrit from the sample. If anything, this would be likely to positively bias our estimates of SFC's impacts on PSI.

5   Empirical Results

Does SFC impact psychosocial and labor force outcomes?

The SFC program had two main objectives. First, Mercy Corps saw the program as a way to improve the psychosocial well-being and resilience of vulnerable youth. Second, Mercy Corps believed that psychosocial improvements would lead to greater workforce "preparedness" and positively impact labor-related outcomes. Without a strong theoretical foundation for why different dimensions of either psychosocial or labor force measures should be preferred over others and to reduce the number of hypotheses being tested, we focus on the aggregate psychosocial index (PSI) and labor force index (LFI) detailed in the previous section.[16]

---

[15]Stratification of treatment by community alleviates the concern that inconsistency in time between the program and the follow-up survey bias estimation of the program's impact. These differences in timing prevent us from making comparisons of treatment effects across communities.

[16]We follow Olken (2015) and aggregate multiple measures into aggregate indices to reduce the number of hypotheses being tested along with the risks associated with multiple hypothesis testing.

We estimate the direct effects of the program on psychosocial measures and labor outcomes using the following ANCOVA regression specification[17]:

$$Y_{i,t} = \beta_0 + \beta_1 SFC_i + Y_{i,t-1} + \lambda \mathbf{X}_i + \delta_c + \epsilon_i \qquad (1)$$

Where $Y_{i,t}$ is an outcome of interest for individual, i, measured at the endline in time, t, SFC_i is an indicator for whether an individual was assigned to the sports for change program, and $Y_{i,t-1}$ is an individual's baseline level of the outcome of interest. $\mathbf{X}_i$ is a set of time-invariant covariates which include age and age-squared as well as dummies for female and highest grade level attained. We include a set of community fixed effects, δ_c, and use robust standard errors to adjust for heteroskedasticity of the error term.

As discussed in Section 4, take up of the SFC program among the treatment group was high, though not universal. We therefore interpret our estimates of β_1 as the average causal effect of the program for those assigned to the treatment group, i.e. the intention to treat estimate. For our main effects, we also present the treatment on the treated estimate by conducting two stage least squares, with assignment to treatment as an instrument for having ever attended an SFC session. 73% of those assigned to an SFC group attended at least one SFC session whereas no one in the control group ever attended an SFC session. Unsurprisingly, this instrument is highly significant, with a T-stat of 53.6 (see Appendix Table A.3).

Table 5 shows the program's main impacts. The intention to treat estimates are shown in columns (1) and (2) for PSI and LFI, respectively. For PSI, we see a small, negative point estimate with a 95% confidence interval that includes effects between -0.098 and 0.07 standard deviations. However, the LFI impacts are positive and statistically significant. Column (2) shows a point estimate of 0.115 standard deviations (p=0.011) with a

---

[17]With random assignment, two survey periods, and low autocorrelation of our outcome variables, we follow McKenzie (2012) and use ANCOVA as our preferred specification for greatest statistical precision.

confidence interval from 0.025 to 0.205. Testing for two outcomes, we perform a sharpened false discovery rate adjustment following Anderson (2008). Adjusted q-values are reported in brackets beneath the standard errors. The effects of SFC on the labor force index in column (2) retains statistical significance with a q-value of 0.024.

Effects of SFC on the psychosocial sub-indices are shown in Appendix Table A.4. Consistent with the effects on the aggregate measure, they are not encouraging. The estimated treatment effects for the different psychosocial sub-indices have different signs and the only outcome showing marginal significance, risky behaviors, has a negative point estimate.[18] By contrast, Appendix Table A.5 shows effects of SFC on the two sub-components of the labor force index, suggesting 11-12% increases in hours of labor supply and weekly earnings that both survive multiple hypothesis testing (q-value = 0.072).[19]

Columns (3) and (4) of Table 5 show the treatment on the treated effect by using random assignment to an SFC group as an instrument for having ever attended an SFC session in a two-stage least squares estimation. As expected given partial uptake of the treatment, magnitudes of both estimates increase. For LFI, the treatment on the treated estimation suggests an average increase of 0.161 standard deviations for those who attended any SFC sessions, with a 95% confidence interval between 0.039 and 0.283.

Does SFC improve resilience?

The SFC program was also motivated as a way to improve participants' resilience to negative events. To explore this, we look at heterogeneous program effects in the presence of negative life events.

---

[18]Appendix Table A.8 shows treatment effects on all sub-components of the psychosocial index sub-indices. Unsurprisingly, none of them retain statistical significance after correcting for multiple hypothesis testing within each sub-index.

[19]Appendix Table A.7 shows robustness of the main labor force effects to inclusion of noisier 90-day income recall. Because 50% of randomly selected respondents in the SFC treatment group were also invited to participate in the CFW program, we test whether effects attributed to SFC could have been driven by CFW instead. Appendix Table A.6 shows that, point estimates of effects for LFI are larger for those who *were not* in the CFW treatment group.

Our specification for measuring heterogeneous treatment effects is the following:

$$Y_{i,t} = \beta_0 + \beta_1 SFC_i + \beta_2(SFC_i \times Het_i) + \beta_3 Het_i + Y_{i,t-1} + \lambda \mathbf{X}_i + \delta_c + \epsilon_i \qquad (2)$$

Where Het_i is a measure of individual level heterogeneity, such as exposure to negative life events. The primary coefficient of interest is the estimate of β_2, that tests whether program responses differ by this dimension of heterogeneity.[20]

In columns (5) and (6) of Table 5 we test whether the program has additional benefits for those who recently experienced negative life events. Due to the Ebola outbreak, all study participants undoubtedly experienced a meaningful disruption to their daily lives while many likely experienced considerable adversity.[21] Thus the main results on PSI may already suggest limited impacts on resiliency. The heterogeneity analysis allows us to further explore resiliency in the presence of additional household trauma. First, we note a strong negative correlation between the life event index and psychosocial outcomes. This further increases the credibility of these measures and raises our confidence that our psychosocial measures are not mere noise. However, the coefficient of interest in this specification is the interaction term between SFC and negative life events. For PSI, greater resilience from the program would imply a positive coefficient on the interaction. While positive, our estimate for this term is close to zero with a 95% confidence interval from -0.074 to 0.098. We do not, therefore, find evidence of improved psychosocial resilience in the presence of negative life events resulting from the program.

Column (6) of Table 5 again looks at program heterogeneity in the presence of negative life events for the labor force index. However, these predictions are less clear than they were for psychosocial outcomes. Greater resiliency in the presence of negative life

---

[20]This approach could be problematic if the SFC program directly affected probability of participants experiencing negative life events. Appendix Table A.9 looks at the incidence of negative life events in the sample, split by treatment status, and does not find significant differences.

[21]Due to sensitivity of the issue at the time of the follow-up, we did not ask directly about respondents or family members being impacted by Ebola. However, some effects of Ebola were likely captured in the life event index in the form of sick or deceased family members.

events may be reflected in increased labor force participation to cope with these shocks, or it may conversely be reflected in a weaker response if the program increased resiliency by reducing respondents' vulnerability to shocks. The regression results in column (6) suggest that life events are associated with increased labor force participation. However, the interaction term is negative, thus muting this response, with a confidence interval that contains zero. Without a clear prediction and without significant effects we do not consider this to be evidence either for or against greater labor force resiliency.

SFC was motivated by a belief that the program would improve participants' psychosocial well-being and that this would, in turn, improve participants' ability to participate in the labor force. While we do find positive effects on the latter outcome, we do not see evidence of positive impacts on psychosocial measures or improved resiliency. These results suggest that the effects of SFC on labor force outcomes may not be conditional on improvements to psychosocial well-being.

Are treatment effects concentrated among certain subgroups?

In this section, we present a secondary analysis to better understand *who* benefited most from the program and whether impacts are reflected in informative and sensible patterns of heterogeneity. To explore this, we check for heterogeneous treatment effects across a number of different dimensions: gender, age, education, and previous vocational training. Each of these dimensions can be sensibly divided into a group which is economically advantaged (male, older, better educated, or with training) or disadvantaged. We additionally use our core set of covariates to predict labor market outcomes to see if those with high or low predicted labor force outcomes (above or below the median) respond differently to the treatment.[22]

Figure 2 plots the estimated LFI treatment effects from separately estimated ANCOVA

---

[22]Due to low autocorrelation in our main outcome measures (0.107 for LFI and 0.202 for PSI), we use data from the control group at follow-up to generate predicted endline values in the full sample.

regressions for different dimensions of heterogeneity. For each pair of estimates, we find that impacts of SFC on LFI are larger for the more disadvantaged group. We also utilize Equation 2 where Het_i is a dimension which we test for heterogeneous treatment effects. These results are shown in regression form in Panel (a) of Table 6. Each row shows coefficients from a separately estimated version of equation (2) with the main effect of SFC participation, an interaction term, and the dimension of heterogeneity (listed in the first column). The first row of Panel (a) shows a positive point estimate of the SFC program of 0.064 standard deviations, with standard errors of 0.066. Column (2) shows an interaction term of 0.098 that is not significantly different from zero. However, the overall effect for females (SFC + SFCxHet in the table) is 0.162 standard deviations, with a p-value of 0.01 reported in column (5). Although the interaction terms testing the difference between these groups are not significant, we observe that all five of the disadvantaged groups experience an overall positive treatment effect with significance levels at or below 5%.[23]

The final pair of results in Figure 2 and row (5) of Panel (a) in Table 6 show that those with worse predicted labor force outcomes have significantly larger treatment effects than those expected to be doing better in the absence of the program. People predicted to have low labor force outcomes have a treatment effect that is 0.21 standard deviations bigger than those predicted to have higher labor force outcomes in the absence of the program (p=0.02). Given a lack of clear theoretical predictions motivating who we should have expected to have greater or lesser responses to the program, we consider this analysis to be exploratory. However, consistent patterns across multiple dimensions of heterogeneity are noteworthy and may provide a starting point for future research in a limited literature.

We perform a similar set of analyses to look for parallel patterns of program impact on psychosocial outcomes in Table 6, Panel (b). We do not find meaningful positive treatment effects on PSI for any of our subgroups across these dimensions of heterogeneity. Outcomes for those with high predicted PSI may, in fact, be negative, while those with

---

[23]As this secondary analysis is exploratory in nature, we do not correct these p-values for multiple comparisons (following Olken (2015)).

worse predicted outcomes have a positive point estimate that is indistinguishable from zero. We also see slightly worse PSI outcomes (though not statistically significant) for those with worse predicted labor force outcomes.

Overall, we take these patterns as additional evidence that the SFC program did not induce positive changes in psychosocial outcomes among easily identifiable subgroups. That we do not see effects on PSI for those groups driving impacts on labor force outcomes further undermines our confidence that labor force benefits came through a psychosocial channel. Ultimately, the mechanisms behind the positive labor force impacts may be more nuanced than a simplified theory of change based on these intermediate indicators of psychosocial well-being.

Does group composition affect program impacts?

A different possibility is that group composition and peer effects play a central role in program effectiveness. After the results of our evaluation were known, Mercy Corps expressed particular concern that random assignment to youth groups may have disrupted the efficacy of the program on psychosocial impacts. Life skills lessons covered many sensitive topics and required trust among group members, sharing personal experiences with the other members. Randomization of registrants into sports groups may have made groups less cohesive with members sharing less similar backgrounds or pre-existing relationships than if they had been permitted to choose their own groups.

We explore whether program impacts differed depending on different measures of group cohesion in Table 7. Baseline data collection included questions about each respondent's social network, allowing us to identify pre–existing social linkages among study participants. Panel (a) shows that presence of a friend in one's sports group significantly increased the likelihood of ever attending one of the sessions as well as total attendance by 9-10%. However, we do not find evidence that presence of a friend improves either psychosocial or labor force outcomes. Panel (b) examines whether individual outcomes

vary by group ethnic diversity. Group diversity is measured by calculating a Herfindahl index, $H_g = \sum_{i=1}^{N} E_i^2$, where E_i is the share of sports group g's members that belong to a particular ethnic group, i.[24] In Panels (c) and (d) we test for heterogeneity by whether a respondent has more (or less) age or age-gender mates in their randomly assigned sports group. While the presence of a larger number of similar age and gender teammates increases program attendance (panel (d)), their presence does not appear to improve program impacts on psychosocial or labor force measures.

We adopt two additional approaches to examine whether peer influence is an important channel for program effects. First, we test whether exposure to "better" peers could have positive impacts on participants. In particular, we test whether groups with greater average baseline levels of either psychosocial measures or labor force measures have differential outcomes. We calculate a "leave-out-mean" baseline value for each individual's assigned sports group (leaving out the individual's own baseline value in this calculation), subtracting community level averages from these group measures. Respondents in the control group are assigned a value of zero and are preserved in this analysis as a reference point to determine if peer effects could be driving the overall effects of the program. With treatment randomly assigned, and group membership randomly assigned conditional on community, we take these group composition variables as exogenous.

We find little evidence that peer influence is driving the average treatment effect of the program on LFI. Table 8 indicates that controlling for the leave–out mean of group PSI or LFI (normalized to be centered on zero) leaves the coefficient on LFI virtually unaffected and similarly precise (0.118, compared to a coefficient of 0.115 in column (2) of table 5). There is, however, evidence that peer composition may explain some heterogeneity about this mean treatment effect. In particular, panel (a) column (4) of Appendix

---

[24]In our regressions, we standardize the Herfindahl scores using the following procedure. First, we simulate 1,000 draws of each community's sports groups and calculate a mean Herfindahl index, $\widehat{H}_c$, for each community, c. Then we compute the deviation from the simulated index for each sport group and divide by the standard deviation of our simulated Herfindahl Index, $\frac{H_g - \widehat{H}_c}{sd(\widehat{H}_c)}$

Table 8 indicates that an additional 0.1 standard deviation average of baseline PSI is associated with a 0.046 standard deviation improvement in LFI, an effect nearly 40% as large as the main effect of the program. At the same time, we do not find evidence that "better" groups in terms of baseline LFI are associated with larger treatment effects. We also do not find evidence that having a group with a larger baseline PSI or LFI is associated with improved PSI outcomes. Absent a theory of why better PSI groups may matter (only) for LFI outcomes when better LFI groups do not, we leave this analysis as a suggestive area for further research.[25] [26]

Ultimately, while we find some suggestive evidence of peer effects influencing labor force outcomes, these analyses do not provide clear insights into the mechanisms of the program's effects on LFI.

6 Summary and Conclusion

Using sports as a method of intervention and vehicle for socio-emotional and psychosocial training has come increasingly into fashion. Sport for Development (SFD) is viewed as a potentially transformative approach to engaging and positively affecting the lives of vulnerable youth. These programs involve millions of participants across the globe and constitute hundreds of millions of dollars of expenditures each year. Despite these high levels of participation and expenditure, there is little existing evidence on the efficacy of these programs and their effect on participants.

Our evaluation focused on the SFD programming developed and implemented in-

---

[25]We also note that these peer effect estimates do not appear to explain the heterogeneity of the main effects from section 5 where those with low predicted labor force outcomes benefit more from the program. In Appendix Table A.10 we see that peer effects are driven by those who are predicted to have *higher* LFI outcomes. While leaving the channel of the program's direct effects on labor force outcomes unexplained, this adds further credence that the effects are coming through the program and not peer effects.

[26]We additionally performed an agnostic test of peer effects following Shue (2013) with two-way clustering for dyadic data following Fafchamps and Gubert (2007). Our test reveals that group members are, if anything, more dissimilar in their outcomes following intervention than they are to other study participants. This test provides further that positive peer effects are not an important driver of the evaluation results. These results are shown in Appendix Table 11.

ternationally by Mercy Corps, one of the global leaders in this space. In this context, we found that their Sports for Change program exerted limited impacts on psychosocial outcomes, but did increase labor force engagement a year after the intervention by a statistically significant 0.12 standard deviations. We note that it is plausible that similar programs may generate different effects on these outcomes if they use a different method of selecting and training coaches. For example, professional counsellors, trained therapists, or job placement experts may be expected to result in different impacts on psychosocial and labor force outcomes if they serve as coaches, a possibility that we cannot speak to in this paper.

Given this lack of effects on psychosocial measures, the motivating theory of change for SFD does not appear to have been the mechanism driving improved labor market outcomes in this setting. While we were ultimately unable to isolate these mechanisms, heterogeneity analysis suggests that more disadvantaged groups (women, less educated, young, and those without vocational training) benefited most from the program.

Ultimately, this evaluation provides evidence of positive impacts of an SFD program on labor force outcomes. Given the scarcity of positive findings on active labor market programs in developing countries in general and the extent to which SFD programs are not precisely targeted at boosting labor market outcomes, these results are notable in their precision and magnitude (McKenzie 2017). However, the strength of these results is tempered by our inability to identify the mechanism through which this program's impact works. Given the pervasiveness and scale of resources devoted to SFD programs, we feel that further research should be done in order to deepen the pool of evidence on sport for development programs.

References

Adair, Daryl. 2014. *Global Sport-for-Development: Critical Perspectives*. Palgrave Macmillan UK.

Adoho, Franck, Shubha Chakravarty, Dala T. Korkoyah, Mattias K. A. Lundberg, and Afia Tasneem. 2014. "The Impact of an Adolescent Girls Employment Program: The EPAG Project in Liberia." *SSRN Scholarly Paper*, (ID 2420245).

Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association*, 103(484): 1481–1495.

Bendavid, Eran, Trygve Ottersen, Liu Peilong, Rachel Nugent, Nancy Padian, John-Arne Rottingen, and Marco Schaferhoff. 2017. "Development Assistance for Health." *Disease Control Priorities: Improving Health and Reducing Poverty. 3rd edition*. Ed. by D.T. Jamison, H. Gelband, and S. Horton. The International Bank for Reconstruction and Development / The World Bank. Chap. 16, 299–314.

Berlin, Richard A., Aaron Dworkin, Ned Eames, Arn Menconi, and Daniel F. Perkins. 2007. "Examples of Sports-Based Youth Development Programs." *New Directions for Student Leadership*, 2007(115): 85–106.

Blattman, Christopher, Julian C. Jamison, and Margaret Sheridan. 2017. "Reducing Crime and Violence: Experimental Evidence from Cognitive Behavioral Therapy in Liberia". *American Economic Review*, 107(4): 1165–1206.

Burnett, Cora. 2009. "Engaging Sport-for-Development for Social Impact in the South African Context." *Sport in Society: Cultures, Commerce, Media, Politics*, 12(9): 1192–1205.

Calero, Carla and Sandra V. Rozo. 2016. "The Effects of Youth Training on Risk Behavior: The Role of Non-Cognitive Skills." *IZA Journal of Labor & Development*, 5(1): 12.

Cantril, Hadley. 1966. *The Pattern of Human Concerns*. New Brunswick, N.J.: Rutgers University Press.

Coalter, Fred. 2009. "Sport-in-Development: Accountability or Development?" *Sport and International Development*. Ed. by Roger Levermore and Aaron Beacom. Global Culture and Sport. London: Palgrave Macmillan UK. Chap. 3, 55–75.

Coalter, Fred. 2010. "The Politics of Sport-for-Development: Limited Focus Programmes and Broad Gauge Problems?" *International Review for the Sociology of Sport*, 45(3): 295–314.

Deaton, Angus. 2008. "Income, Health, and Well-Being around the World: Evidence from the Gallup World Poll." *Journal of Economic Perspectives*, 22(2): 53–72.

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): 380–392.

Donnelly, Peter, Simon Darnell, S Wells, and J Coakley. 2007. *The Use of Sport to Foster Child and Youth Development and Education.* Toronto.

Dunmore, Charlie. 2013. "Liberian President Says Youth Unemployment a Threat to Peace". *Reuters, 25 November*.

Fafchamps, Marcel and Flore Gubert. 2007. "The formation of risk sharing networks". *Journal of Development Economics*, 83(2): 326–350.

Garlick, Robert, Kate Orkin, and Simon Quinn. Forthcoming. "Call Me Maybe: Experimental Evidence on Using Mobile Phones to Survey Microenterprises." *World Bank Economic Review*.

Groh, Matthew, Nandini Krishnan, David McKenzie, and Tara Vishwanath. 2016. "The Impact of Soft Skills Training on Female Youth Employment: Evidence from a Randomized Experiment in Jordan." *IZA Journal of Labor & Development*, 5(1): 9.

Hartmann, Douglas and Brooks Depro. 2006. "Rethinking Sports-Based Community Crime Prevention: A Preliminary Analysis of the Relationship between Midnight Basketball and Urban Crime Rates." *Journal of Sport and Social Issues*, 30(2): 180–196.

Hartmann, Douglas and Christina Kwauk. 2011. "Sport and Development: An Overview, Critique, and Reconstruction." *Journal of Sport & Social Issues*, 35(3): 284–305.

Heckman, James J. and Tim Kautz. 2012. "Hard Evidence on Soft Skills." *Labour Economics*, 19(4): 451–464.

Henley, Robert, Ivo Schweizer, Francesco de Gara, and Stefan Vetter. 2007. "How Psychosocial Sport & Play Programs Help Youth Manage Adversity: A Review of What We Know & What We Should Research." *International Journal of Psychosocial Rehabilitation*, 12(1): 51–58.

Holt, Nicholas L. and Martin I. Jones. 2007. "10 Future Directions for Positive Youth Development and Sport Research." *Positive Youth Development through Sport*. Ed. by Nicholas L. Holt. Routlidge. Chap. 10, 122–132.

Ibarraran, Pablo, Laura Ripani, Bibiana Taboada, Juan Miguel Villa, and Brigida Garcia. 2014. "Life Skills, Employability and Training for Disadvantaged Youth: Evidence from a Randomized Evaluation Design." *IZA Journal of Labor & Development*, 3(1): 10.

Jeanes, Ruth. 2013. "Educating through Sport? Examining HIV/AIDS Education and Sport-for-Development through the Perspectives of Zambian Young People." *Sport, Education and Society*, 18(3): 388–406.

Kautz, Tim, James J Heckman, Ron Diris, Bas ter Weel, and Lex Borghans. 2014. *Fostering and Measuring Skills: Improving Cognitive and Non-Cognitive Skills to Promote Lifetime Success*. Working Paper 20749. National Bureau of Economic Research.

Kidd, Bruce. 2008. "A New Social Movement: Sport for Development and Peace." *Sport in Society*, 11(4): 370–380.

Lowe, Matt. 2018. "Types of Contact: A Field Experiment on Collaborative and Adversarial Caste Integration." *Mimeo*.

McKenzie, David. 2012. "Beyond Baseline and Follow-up: The Case for More T in Experiments." *Journal of Development Economics*, 99(2): 210–221.

McKenzie, David. 2017. "How Effective Are Active Labor Market Policies in Developing Countries? A Critical Review of Recent Evidence." *The World Bank Research Observer*, 32(2): 127–154.

Morris, Leesa, Jo Sallybanks, Katie Willis, and Toni Makkai. 2004. "Sport, Physical Activity and Antisocial Behaviour in Youth." *Youth Studies*, 4(1): 47–52.

Mousa, Salma. 2018. "Overcoming the Trust Deficit: Intergroup Contact and Associational Life in Post-ISIS Iraq." *Mimeo*.

Olken, Benjamin A. 2015. "Promises and perils of pre-analysis plans." *Journal of Economic Perspectives*, 29(3): 61–80.

Organization for Economic Cooperation and Development. 2019. *Development Aid at a Glance.* https://www.oecd.org/dac/financing-sustainable-development/development-finance-data/World-Development-Aid-at-a-Glance-2019.pdf.

Panter-Brick, Catherine, Rana Dajani, Mark Eggerman, Sabrina Hermosilla, Amelia Sancilio, and Alastair Ager. 2018. "Insecurity, Distress and Mental Health: Experimental and Randomized Controlled Trials of a Psychosocial Intervention for Youth Affected by the Syrian Crisis." *Journal of Child Psychology and Psychiatry*, 59(5): 523–541.

Perkins, Daniel F and Gil G Noam. 2007. "Characteristics of Sports-Based Youth Development Programs." *New Directions for Student Leadership*, 2007(115): 75–84.

Petitpas, Albert J, Allen E Cornelius, Judy L Van Raalte, Tiffany Jones, et al. 2005. "A Framework for Planning Youth Sport Programs That Foster Psychosocial Development." *The sport psychologist*, 19(1): 63–80.

Raine, Adrian, Kenneth Dodge, Rolf Loeber, Lisa Gatzke-Kopp, Don Lynam, Chandra Reynolds, Magda Stouthamer-Loeber, and Jianghong Liu. 2006. "The Reactive–Proactive Aggression Questionnaire: Differential Correlates of Reactive and Proactive Aggression in Adolescent Boys." *Aggressive Behavior*, 32(2): 159–171.

Rookwood, Joel. 2008. "Soccer for Peace and Social Development." *Peace Review*, 20(4): 471–479.

Rosenberg, Morris. 1965. *Society and the Adolescent Self-Image*. Princeton, N.J.: Princeton University Press.

Rotter, Julian B. 1966. "Generalized Expectancies for Internal versus External Control of Reinforcement." *Psychological monographs: General and applied*, 80(1): 1–28.

Shue, Kelly. 2013. "Executive networks and firm policies: Evidence from the random assignment of MBA peers". *The Review of Financial Studies*, 26(6): 1401–1442.

Svensson, PG and H Woods. 2017. "A Systematic Overview of Sport for Development and Peace Organisations". *Journal of Sport for Development*, 5(9): 36–48.

Underwood, Emily. 2018. "Lessons in Resilience." *Science*, 359(6379): 976–979.

Tables

Table 1: Registration and Baseline Survey Dates and Participants

Community	Registration	Baseline Interviews	SFC		Control		Total
			Male	Female	Male	Female	
West Point	24 Jul 2013	29 Jul – 8 Aug 2013	90	70	91	69	320
New Kru Town	7 Sep 2013	11 Sep – 24 Sep 2013	116	124	117	123	400
Peace Island	31 Oct 2013	3 Nov – 8 Nov 2013	39	41	41	39	160
Buzzy Quarter	7 Nov 2013	9 Nov – 16 Nov 2013	38	42	40	40	160
Clara Town	18 Nov 2013	22 Nov – 5 Dec 2013	76	84	78	82	320
Dry Rice Market	12 Dec 2013	15 Dec – 19 Dec 2013	40	40	40	40	160
Banjor	11 Jan 2014	13 Jan – 16 Jan 2014	38	42	40	40	160
Chicken Soup Factory	18 Jan 2014	23 Jan – 30 Jan 2014	68	92	67	93	320
Logan Town	1 Feb 2014	4 Feb – 12 Feb 2014	69	91	65	95	320
Total			574	626	579	621	2400

Table 2: Baseline Balance by Treatment Status

	Control	Treatment	P-value of Diff
Demographics and household characteristics:			
Female	0.518	0.522	0.85
Age	20.926	20.798	0.26
Head of Household	0.131	0.143	0.42
Household Size	6.762	6.598	0.26
Mother known to be living	0.868	0.870	0.91
Father known to be living	0.713	0.717	0.84
Has at least one child	0.454	0.432	0.28
Christian	0.876	0.863	0.37
Muslim	0.111	0.125	0.29
Matched connections in sample	2.390	2.419	0.72
Education and cognitive ability:			
Completed primary school	0.835	0.836	0.95
Completed secondary school	0.276	0.265	0.57
Highest Grade Completed	11.389	11.350	0.81
Numeracy (0–10)	5.951	5.784	0.04**
Digits Forward (0–8)	5.176	5.124	0.42
Digits Backward (0–8)	1.981	1.924	0.24
First Word Recall (0–10)	3.059	3.072	0.91
Second Word Recall (0–10)	2.578	2.558	0.85
Ravens Score (0–3)	1.753	1.726	0.51
Risk Aversion (0–6)	3.785	3.758	0.79
Psychosocial Measures:			
Subj. Welfare, today, (Standardized)	-0.011	0.011	0.59
Self-esteem Index	22.977	22.495	0.00***
Locus of Control Index	24.089	24.031	0.63
Aggression Index	2.603	2.546	0.63
Risky behavior index	1.501	1.468	0.67
Depression, Anxiety and Stress Score (DASS21)	21.987	23.282	0.02**
Psychosocial Index (Standardized)	0.000	-0.065	0.12
Labor Force and Financial Measures:			
Labor Force Index (Standardized)	-0.000	0.021	0.62
Working Last Week	0.431	0.436	0.83
Hours worked last week	11.546	12.344	0.33
Hours worked last week (if positive)	27.495	29.016	0.27
IHST (Hours worked last week)	1.513	1.563	0.53
Income earned last week	6.501	6.647	0.80
Income earned last week (if positive)	16.342	16.999	0.57
IHST (Income earned last week)	1.190	1.180	0.88
Income earned last three months	48.060	57.840	0.01**
Income earned last three months (if positive)	80.078	96.591	0.00***
IHST (Income earned last three months)	2.625	2.738	0.25
N	1197	1198	

Notes: Top 1% of values for hours worked, income earned last week, and income earned last three months are trimmed and set to missing. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Baseline Labor Force Index (LFI) Correlations with Psychosocial Measures

	(1) LFI	(2) LFI	(3) LFI	(4) LFI	(5) LFI	(6) LFI
Welfare Today (WF)	0.068*** (0.021)					
Self-Esteem Index (SE)		0.109*** (0.020)				
Locus of Control Index (LOC)			0.056*** (0.020)			
Aggressive Behaviors Index (Agro)				0.031 (0.021)		
Risky Behaviors Index (RB)					0.006 (0.023)	
Psychosocial Index (PSI)						0.116*** (0.021)
N	2310	2310	2310	2310	2310	2310
Mean Y	0.01	0.01	0.01	0.01	0.01	0.01
R2	0.061	0.068	0.060	0.058	0.057	0.069

Notes: All indices coded so that positive values indicate "better" behavior. Psychosocial index is the standardized mean of the five (standardized) sub-indices. Standard errors in parentheses. Top 1% of earnings and hours are trimmed and set to missing to prevent relationships driven by implausibly large outliers. Regressions include female, age, and age^2 covariates as well as fixed effects for educational attainment and community. All variables are coded so that higher values reflect "better" behaviors or attitudes. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Sports for Change Life Skills Sessions: Topics and Skills Addressed

Session	Sports for Change Skills				
	Resilience	Planning	Trust	Self-Esteem	Communication
1: Introduction			X		
2: Me and Others			X	X	X
3: Understanding Emotions	X	X	X	X	X
4: Communication			X		X
5: Relationships	X		X		X
6: Cooperation			X	X	X
7: Believing in Me		X	X		X
8: Conflict and Violence	X	X	X	X	X
9: Collaboration			X		X
10: Motivation	X	X		X	
11: Dealing with Problems	X		X	X	X
12: Making Strategies		X	X	X	
13: Applying SFC Skills in My Life	X	X	X	X	X
14: Planning Graduation Event		X			
15: What Have We Learned?	X	X	X	X	X
16: Review and Closing	X	X	X	X	X

Table 5: Impact of SFC – ANCOVA

	(1) PSI	(2) LFI	(3) PSI	(4) LFI	(5) PSI	(6) LFI
SFC	-0.014	0.115**			-0.016	0.123***
	(0.042)	(0.045)			(0.042)	(0.046)
	[0.581]	[0.024]			[0.634]	[0.034]
Attended SFC			-0.026	0.156***		
			(0.057)	(0.062)		
			[0.581]	[0.024]		
SFC x Life Event Index					0.012	-0.062
					(0.044)	(0.044)
					[0.634]	[0.311]
Life Event Index					-0.059**	0.071**
					(0.030)	(0.030)
Life Event Index (Lagged)					-0.008	0.025
					(0.021)	(0.023)
Dependent Var (Lagged)	0.185***	0.098***	0.099***	0.015	0.184***	0.098***
	(0.021)	(0.024)	(0.027)	(0.024)	(0.023)	(0.024)
N	2055	1947	2055	1947	2055	1947
Mean Y	0.003	0.054	0.003	0.054	0.003	0.054
Control Mean	0.008	-0.002	0.008	-0.002	0.008	-0.003
P-Val	0.734	0.011	0.647	0.012	0.696	0.008
P-Val 2					0.776	0.158
R2	0.077	0.075	0.054	0.071	0.080	0.078

Notes: PSI = Psychosocial Index. LFI=Labor Force Index. FDR sharpened q-values for the main coefficients of interest are calculated following Anderson (2008) and presented in brackets. Q-values for the coefficient on SFC in columns (1) and (2) are calculated with adjustments for two possible outcomes. Similarly q-values are adjusted for the coefficient on any SFC attendance in columns (3) and (4). The regressions on resilience adjust for four possible outcomes in columns (5) and (6) for the main effects of SFC and the interactions with the life event index. All regressions also include controls for age and age-squared, as well as dummies for gender, educational attainment, and community fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Heterogeneous Impacts on PSI and LFI

Panel (a): Labor Force Index	SFC	SFCxHet	Het	Y_L1	P:$\beta1 + \beta2$	Control Group Means $Het = 1$	$Het = 0$
Female	0.063	0.098	-0.253	0.095	0.010	-0.131	0.138
	(0.066)	(0.091)	(0.063)	(0.024)			
Young	0.082	0.061	-0.006	0.096	0.016	-0.146	0.160
	(0.070)	(0.092)	(0.102)	(0.024)			
Low Education	0.069	0.092	0.046	0.093	0.024	-0.111	0.079
	(0.060)	(0.093)	(0.239)	(0.024)			
No Training	0.014	0.160	-0.116	0.096	0.002	-0.069	0.107
	(0.078)	(0.097)	(0.066)	(0.024)			
Predicted Low LFI (0/1)	0.010	0.210	-0.161	0.094	0.000	-0.261	0.255
	(0.067)	(0.091)	(0.090)	(0.024)			
Predicted Low PSI (0/1)	0.073	0.086	-0.112	0.095	0.016	-0.046	0.041
	(0.063)	(0.091)	(0.087)	(0.024)			
Panel (b): Psychosocial Index	SFC	SFCxHet	Het	Y_L1	P:$\beta1 + \beta2$	$Het = 1$	$Het = 0$
Female	-0.051	0.072	-0.068	0.180	0.707	-0.064	0.087
	(0.062)	(0.084)	(0.061)	(0.021)			
Young	-0.034	0.038	0.154	0.180	0.946	0.024	-0.010
	(0.062)	(0.084)	(0.089)	(0.021)			
Low Education	-0.030	0.042	-0.386	0.181	0.861	-0.144	0.101
	(0.055)	(0.086)	(0.202)	(0.021)			
No Training	0.014	-0.042	-0.093	0.180	0.591	-0.049	0.102
	(0.070)	(0.087)	(0.061)	(0.021)			
Predicted Low LFI (0/1)	0.024	-0.073	0.082	0.180	0.406	0.011	0.011
	(0.059)	(0.084)	(0.083)	(0.021)			
Predicted Low PSI (0/1)	-0.100	0.175	-0.093	0.179	0.212	-0.192	0.209
	(0.058)	(0.084)	(0.080)	(0.021)			

Notes: LFI=Labor force index. PSI=Psychosocial index. Each row results from a separately estimated ANCOVA regression with a dimension of heterogeneity listed in the first column. Standard errors for each covariate listed in columns 1-4 are listed in parentheses in the row just below. Predicted LFI and predicted PSI are calculated by separately regressing each outcome on female, age, age^2, no training, and community and educational attainment fixed effects in the control group sample. Using these estimated coefficients, we then split the sample by high or low predicted LFI and PSI measures at the median.

Table 7: Effects of Group Cohesion

Panel (a): Friends	(1) Ever Attend	(2) Total Attendance	(3) PSI	(4) LFI
SFC	0.705***	10.349***	-0.025	0.123**
	(0.020)	(0.283)	(0.040)	(0.046)
SFC x Any Friends	0.079***	1.211***	0.008	-0.015
	(0.028)	(0.421)	(0.067)	(0.068)
Mean Any Friends	0.365	0.365	0.365	0.365
R2	0.598	0.590	0.076	0.086

Panel (b): Group Diversity	(1) Ever Attend	(2) Total Attendance	(3) PSI	(4) LFI
SFC	0.738***	10.818***	-0.006	0.100***
	(0.019)	(0.289)	(0.029)	(0.036)
SFC x Herfindahl SD	0.004	-0.024	0.015	-0.029
	(0.008)	(0.137)	(0.019)	(0.019)
R2	0.5900	0.5818	0.0776	0.0751

Panel (c): Similar Ages	(1) Ever Attend	(2) Total Attendance	(3) PSI	(4) LFI
SFC	0.655***	9.595***	-0.119	0.377**
	(0.073)	(1.107)	(0.165)	(0.140)
# Grp Members w/ Similar Age	0.004	0.064	0.005	-0.013*
	(0.003)	(0.053)	(0.008)	(0.007)
Mean Grp Members w/ Similar Age	19.150	19.150	19.150	19.150
R2	0.5919	0.5837	0.0753	0.0779

Panel (d): Similar Age and Genders	(1) Ever Attend	(2) Total Attendance	(3) PSI	(4) LFI
SFC	0.644***	9.420***	-0.130	0.379**
	(0.072)	(1.097)	(0.169)	(0.141)
# Grp Members w/ Similar Age	-0.005	-0.084	0.004	-0.020*
	(0.006)	(0.088)	(0.011)	(0.011)
# Grp Members w/ Similar Age and Gender	0.018**	0.303**	0.002	0.012
	(0.008)	(0.112)	(0.015)	(0.016)
Mean Grp Members w/ Similar Age	19.150	19.150	19.150	19.150
Mean Grp Members w/ Similar Age and Gender	9.941	9.941	9.941	9.941
R2	0.5992	0.5910	0.0809	0.0871

	(1)	(2)	(3)	(4)
N	2058	2058	2058	1947
Treatment Mean	0.3678	5.4111	-0.0180	0.1078
Control Mean	0.0000	0.0000	-0.0000	-0.0033

Notes: Results from ANCOVA estimation. Herfindahl SD reflects standard deviations of actual group Herfindahl index from simulated community level means values. "Similar Age" is defined as the number of other group members who are similarly either above or below the median age in the sample. "Similar Age and Gender" restricts this measure further by also requiring that the group member is of the same gender as well as age group. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Leave Out Mean Effects

Panel (a): Leave Out Means by Baseline PSI

	(1) Ever Attend	(2) Total Attendance	(3) PSI	(4) LFI
SFC	0.738***	10.850***	-0.019	0.118***
	(0.018)	(0.269)	(0.033)	(0.033)
Group Leave Out Mean	-0.126	-1.371	-0.092	0.461***
	(0.095)	(1.306)	(0.128)	(0.157)
LFI (Lagged)	-0.006	-0.067	0.023	0.094***
	(0.007)	(0.113)	(0.026)	(0.024)
PSI (Lagged)	-0.006	-0.096	0.174***	0.041*
	(0.008)	(0.128)	(0.024)	(0.021)
N	2007	2007	2007	1947
Mean Y	0.3687	5.4210	-0.0119	0.0539
Control Mean	0.0000	0.0000	-0.0000	-0.0033
R2	0.5948	0.5861	0.0667	0.0783

Panel (b): Leave Out Means by Baseline LFI

	(1) Ever	(2) Tot	(3) PSI	(4) LFI
SFC	0.739***	10.881***	-0.024	0.118***
	(0.018)	(0.264)	(0.034)	(0.037)
Group Leave Out Mean	-0.054	-1.052	0.298	-0.144
	(0.099)	(1.483)	(0.238)	(0.261)
LFI (Lagged)	-0.006	-0.071	0.025	0.092***
	(0.007)	(0.111)	(0.026)	(0.023)
PSI (Lagged)	-0.006	-0.096	0.175***	0.040*
	(0.008)	(0.128)	(0.024)	(0.021)
N	2007	2007	2007	1947
Mean Y	0.3687	5.4210	-0.0119	0.0539
Control Mean	0.0000	0.0000	-0.0000	-0.0033
R2	0.5941	0.5858	0.0671	0.0763

Notes: Group leave out mean is calculated as the average baseline PSI or LFI level of group members, excluding oneself. Group leave out means are calculated to be centered on zero for each community, thus showing standardized deviations from community averages. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figures

Figure 1: Sessions Attended by SFC Participants
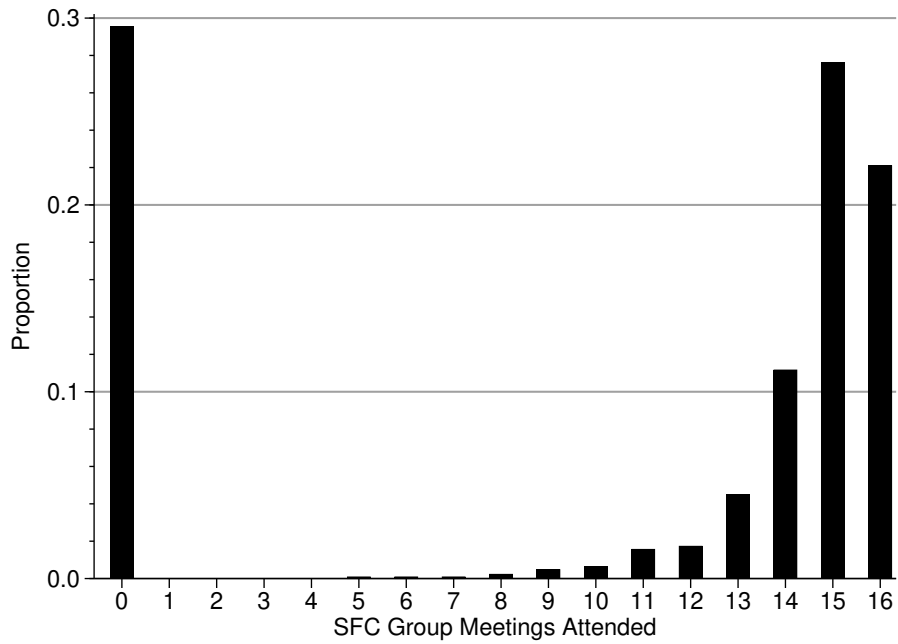
Notes: This figure shows the total number of SFC sessions attended by study participants included in the treatment group. Over 70% attended at least one session while 65% invited to participate attended at least 80% of their group's total number of meetings.

Figure 2: Heterogeneity of SFC Impacts on Labor Force Index
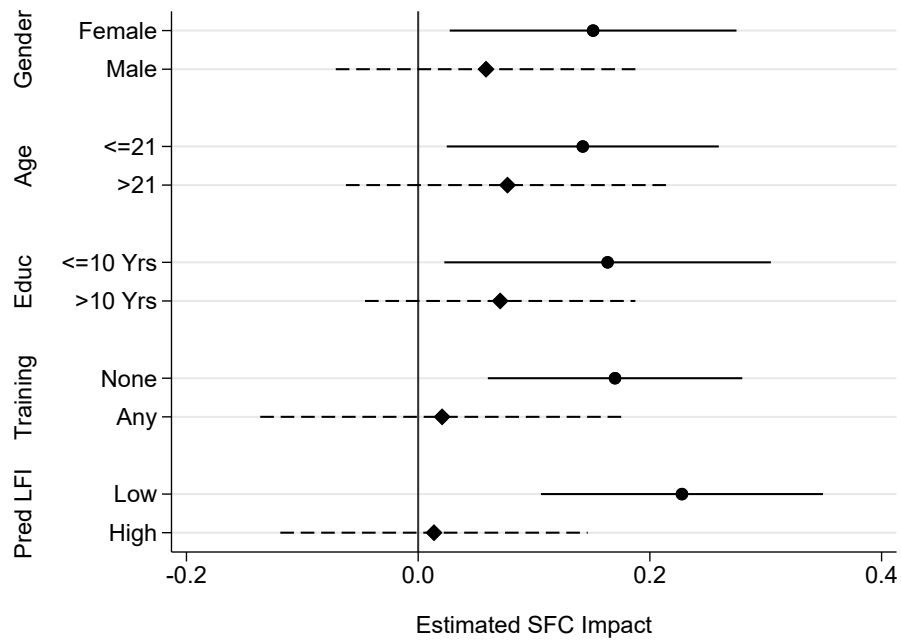
Notes: Figure shows point estimates and 95% confidence intervals of ANCOVA regression of labor force index on program treatment status. Each dimension of heterogeneity splits the sample in two groups as described. "Pred LFI" refers to predicted LFI which is generated with a linear prediction of the labor force index based on predictors including the other dimensions of heterogeneity in the figure along with community fixed effects in the control group. These predictions are then projected onto all study participants and split at the median to indicate those who are above ("High") or below ("Low") the median.

ONLINE APPENDICES

Stay in the Game:

A Randomized Controlled Trial of a Sports and Life Skills Program for Vulnerable Youth in Liberia

Lori Beaman

Northwestern University

Sylvan Herskowitz

International Food Policy Research Institute

Niall Keleher

Stanford University

Jeremy Magruder

University of California, Berkeley

1

Appendix A: Additional Tables and Figures

Table A.1: Test for Selective Attrition

	(1)	(2)	(3)
SFC	-0.001	-0.005	-0.108
	(0.014)	(0.014)	(0.177)
Baseline LFI x SFC			-0.007
			(0.014)
Baseline PSI x SFC			-0.038***
			(0.014)
Female x SFC			0.008
			(0.033)
Age x SFC			0.002
			(0.006)
Head of Household x SFC			-0.025
			(0.041)
Household Size x SFC			-0.003
			(0.004)
Mother Living x SFC			0.028
			(0.041)
Father Living x SFC			-0.004
			(0.031)
Has Children x SFC			-0.067*
			(0.036)
Christian x SFC			0.130
			(0.124)
Muslim x SFC			0.080
			(0.129)
Friends in Sample x SFC			-0.001
			(0.007)
Education Attainment x SFC			-0.004
			(0.004)
Numeracy x SFC			0.002
			(0.008)
N	2395	2333	2333
Mean Y	0.132	0.132	0.132
P-value from F-Test of Joint Significance			0.979
Controls	No	Yes	Yes
R2	0.000	0.034	0.041

Notes: These are results from a regression of a binary indicator of attrition from the study on the treatment and a set of covariates. Column (1) does not include any co-variates. Column (2) includes a set of covariates. Column (3) includes this same set of covariates as well as an interaction term with each covariate and treatment. The p value of the F-test for joint significance of the treatment and its interactions is shown below column (3). Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.2: Labor Outcomes and Jointly Estimated Psychosocial Associations

	(1) Hours	(2) IHST (7dI)	(3) LFI
Welfare Today (WF)	0.679*	0.100***	0.053***
	(0.391)	(0.033)	(0.021)
Self-Esteem Index (SE)	1.417***	0.171***	0.098***
	(0.409)	(0.033)	(0.021)
Locus of Control Index (LOC)	0.706*	0.081**	0.047**
	(0.402)	(0.031)	(0.020)
Aggressive Behaviors Index (Agro)	0.532	0.077**	0.040*
	(0.453)	(0.034)	(0.022)
Risky Behaviors Index (RB)	-0.318	0.018	-0.004
	(0.458)	(0.039)	(0.024)
N	2324	2350	2310
Mean Y	12.02	1.19	0.01
R2	0.051	0.088	0.074

Notes: Top 1% of earnings and hours are trimmed and set to missing to prevent relationships driven by implausibly large outliers. Regressions include female, age, and age^2 covariates as well as fixed effects for educational attainment and community. LFI=Labor Force Index. 7dI=Income over the last seven days. All variables are coded so that higher values reflect "better" behaviors or attitudes. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.3: SFC Assignment and SFC Attendance

	(1) Ever Attend	(2) # Sessions Attended
SFC	0.736***	10.830***
	(0.014)	(0.206)
Female	0.009	0.203
	(0.015)	(0.221)
Age	0.015	0.123
	(0.036)	(0.536)
Age-Squared	-0.000	-0.002
	(0.001)	(0.013)
N	2058	2058
Mean Y	0.367	5.399
Control Mean	0.000	0.000
T-Stat for SFC	53.621	52.679
R2	0.590	0.582

Notes: Column (1) is the first stage of treatment on the treated estimation. Regressions include educational attainment and community fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.4: SFC, Psychosocial Outcomes, and Resiliency

Panel (a): Impact of SFC on Psychosocial Outcome

	(1) WF	(2) SE	(3) LOC	(4) Agro	(5) RB	(6) PSI
SFC	0.024	0.035	-0.009	-0.076*	-0.004	-0.014
	(0.044)	(0.042)	(0.042)	(0.043)	(0.042)	(0.042)
	[1.000]	[1.000]	[1.000]	[0.589]	[1.000]	–
Dependent Var (Lagged)	0.131***	0.214***	0.109***	0.196***	0.252***	0.185***
	(0.024)	(0.021)	(0.022)	(0.029)	(0.032)	(0.021)
N	2055	2055	2055	2055	2055	2055
R2	0.058	0.073	0.044	0.066	0.140	0.077

Panel (b): Impact of SFC on Resiliency – Psychosocial Outcome

	(1) WF	(2) SE	(3) LOC	(4) Agro	(5) RB	(6) PSI
SFC	0.024	0.037	-0.006	-0.073*	-0.017	-0.016
	(0.045)	(0.043)	(0.042)	(0.042)	(0.042)	(0.042)
	[1.000]	[1.000]	[1.000]	[0.574]	[1.000]	–
SFC x Life Event Index	0.006	-0.031	0.002	-0.053	0.098**	0.012
	(0.044)	(0.043)	(0.044)	(0.043)	(0.045)	(0.044)
	[1.000]	[1.000]	[1.000]	[1.000]	[0.389]	–
Life Event Index	-0.012	-0.020	-0.030	0.062***	-0.020	-0.008
	(0.021)	(0.022)	(0.021)	(0.022)	(0.020)	(0.021)
Life Event Index (Lagged)	-0.012	-0.020	-0.030	0.062***	-0.020	-0.008
	(0.021)	(0.022)	(0.021)	(0.022)	(0.020)	(0.021)
Dependent Var (Lagged)	0.130***	0.203***	0.107***	0.181***	0.253***	0.184***
	(0.024)	(0.021)	(0.021)	(0.029)	(0.031)	(0.021)
N	2055	2055	2055	2055	2055	2055
R2	0.058	0.100	0.046	0.075	0.143	0.080

Notes: Outcomes standardized and coded so that positive reflects "better" behavior. WF=Welfare, SE=Self-Esteem, LOC=Locus of Change, Agro=Aggression, RB=Risky Behavior, PSI=Psychosocial Index. Life Event Index is coded with higher values indicating the occurrence of more negative life events. Standard errors in parentheses. Q-values are reported in brackets following Anderson (2008). Panel (a) calculates Q-values adjusting for five simultaneous hypotheses with the P-values from the SFC coefficient. Panel (b) calculates Q-values adjusting for ten simultaneous hypotheses with the P-values from the SFC and Life Event Index x SFC coefficients. All regressions also include controls for age and age-squared, as well as dummies for gender, educational attainment, and community fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.5: SFC, Labor Outcomes, and Resiliency

Panel (a): SFC Impacts on Labor Outcomes

	(1) Hours	(2) IHST (7dI)	(3) LFI
SFC	1.393*	0.119*	0.115**
	(0.724)	(0.065)	(0.045)
	[0.072]	[0.072]	–
Dependent Var (Lagged)	0.069***	0.076***	0.098***
	(0.020)	(0.021)	(0.024)
N	1978	2014	1947
Mean Y	12.483	2.188	0.054
Control Mean	11.734	2.133	-0.003
R2	0.045	0.071	0.075

Panel (b): SFC Impacts on Labor Outcomes with Resiliency

	(1) Hours	(2) IHST (7dI)	(3) LFI
SFC	1.554**	0.125*	0.123***
	(0.736)	(0.066)	(0.046)
	[0.132]	[0.132]	–
SFC x Life Event Index	-1.098	-0.072	-0.062
	(0.708)	(0.064)	(0.044)
	[0.132]	[0.152]	–
Life Event Index	1.221**	0.068	0.071**
	(0.490)	(0.046)	(0.030)
Life Event Index (Lagged)	0.267	0.050	0.025
	(0.357)	(0.033)	(0.023)
Dependent Var (Lagged)	0.067***	0.077***	0.097***
	(0.020)	(0.021)	(0.023)
N	1978	2014	1947
Control Mean	11.734	2.133	-0.002
R2	0.048	0.074	0.078

Notes: LFI=Labor Force Index. 7dI=Income over last 7 days. Life Event Index is coded with higher values indicating more bad events. Standard errors in parentheses. Q-values are reported in brackets following Anderson (2008). Panel (a) calculates Q-values adjusting for two simultaneous hypotheses with the P-values from the SFC coefficient. Panel (b) calculates Q-values adjusting for four simultaneous hypotheses with the P-values from the SFC and Life Event Index x SFC coefficients. All regressions also include controls for age and age-squared, as well as dummies for gender, educational attainment, and community fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.6: Robustness of SFC Impacts to CFW Interactions

	(1) PSI	(2) LFI
SFC	-0.036	0.142**
	(0.052)	(0.055)
SFC x CFW	0.043	-0.054
	(0.060)	(0.066)
Dependent Var - Lagged	0.186***	0.098***
	(0.021)	(0.023)
N	2055	1947
Mean Y	0.003	0.054
Control Mean	0.000	-0.002
R2	0.078	0.075

Notes: PSI=Psychosocial Index. LFI=Labor Force Index This table shows robustness of results to accounting for the cross-randomized Cash for Work (CFW) treatment. Respondents who were only assigned to CFW (and not SFC) were not included in the endline due to budget limitations. Standard errors in parentheses. All regressions also include controls for age and age-squared, as well as dummies for gender, educational attainment, and community fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.7: SFC, Financial Outcomes, and Resiliency with Three Month Income

Panel (a): SFC Impacts on Financial Outcomes

	(1) Hours	(2) IHST (7dI)	(3) IHST (3mI)	(4) LFI
SFC	1.393*	0.119*	0.071	0.103**
	(0.724)	(0.065)	(0.074)	(0.045)
Dependent Var (Lagged)	0.069***	0.076***	0.076***	0.116***
	(0.020)	(0.021)	(0.016)	(0.023)
N	1978	2014	2014	1925
Control Mean	11.734	2.133	3.774	-0.004
R2	0.045	0.071	0.063	0.085

Panel (b): SFC Impacts on Financial Outcomes with Resiliency

	(1) Hours	(2) IHST (7dI)	(3) IHST (3mI)	(4) LFI
SFC	1.554**	0.125*	0.077	0.111**
	(0.736)	(0.066)	(0.076)	(0.046)
SFC x Life Event Index	-1.098	-0.072	-0.091	-0.068
	(0.708)	(0.064)	(0.072)	(0.043)
Life Event Index	1.221**	0.068	0.085*	0.070**
	(0.490)	(0.046)	(0.051)	(0.030)
Live Event Index (Lagged)	0.267	0.050	0.082**	0.033
	(0.357)	(0.033)	(0.037)	(0.022)
Dependent Var (Lagged)	0.067***	0.077***	0.078***	0.116***
	(0.020)	(0.021)	(0.016)	(0.023)
N	1978	2014	2014	1925
Control Mean	11.734	2.133	3.774	-0.004
R2	0.048	0.074	0.067	0.089

Notes: LFI=Labor Force Index. 7dI=Income over last 7 days. 3mI=Income over last 3 months. Life Event Index is coded with higher values indicating more bad events. Standard errors in parentheses. Standard errors in parentheses. All regressions also include controls for age and age-squared, as well as dummies for gender, educational attainment, and community fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.8: Impact of SFC on Psychosocial Sub-Indices and Underlying Questions

	(1) SFC	(2) Std Err	(3) P-Val	(4) Q-Val	(5) Control Mean
Psychosocial Index (PSI)	-0.0206	0.0438	0.6375	–	0.0000
1) Welfare today	0.0242	0.0438	0.5804	1.000	0.0000
2) Locus of Control Index	-0.0165	0.0441	0.7086	1.000	0.0000
State your response to the following statements. 1=Agree, 2=Weakly Agree, 3=Weakly Disagree, 4=Strongly Disagree					
Your choices affect your future.	-0.0570	0.0359	0.1131	1.000	1.4730
You have limited control over your life.	-0.0393	0.0481	0.4139	1.000	1.9846
Business success is the result of luck.	0.0235	0.0526	0.6543	1.000	2.2312
Trying hard can improve your life.	-0.0008	0.0199	0.9673	1.000	1.1272
When you make plans, you can usually make them work.	0.0075	0.0359	0.8349	1.000	1.4923
Bad things that happen to people are due to bad luck.	-0.0043	0.0537	0.9367	1.000	2.9267
People who live on the streets are responsible for their situation.	0.0548	0.0546	0.3160	1.000	2.6888
Success comes from hard work.	0.0149	0.0290	0.6069	1.000	1.2437
3) Self-Esteem Index	0.0299	0.0432	0.4890	1.000	0.0000
State your response to the following statements. 1=Agree, 2=Weakly Agree, 3=Weakly Disagree, 4=Strongly Disagree					
You are satisfied with yourself.	0.0050	0.0468	0.9156	1.000	1.7303
You feel useless sometimes	0.0215	0.0549	0.6956	1.000	2.5400
Everything you try to do, you fail.	0.0074	0.0531	0.8888	1.000	2.4210
You do not have enough respect for yourself.	-0.0585	0.0465	0.2089	0.807	3.4137
You feel that you are at least as good a person as most people.	-0.1066	0.0455	0.0193	0.183	1.7367
You think that you are a good person but doing nothing.	0.0892	0.0560	0.1116	0.641	2.1272
You can do business as well as most others.	0.0304	0.0369	0.4104	1.000	1.4133
You feel ashamed of how your life is going.	-0.0093	0.0547	0.8657	1.000	2.2563
4) Aggression Index	-0.0757	0.0425	0.0748	0.589	0.0000
Over the last year have your friends … (1=often, 2=sometimes, 3=rarely, 4=never)					
…had disputes with leaders/elders.	0.0193	0.0351	0.5818	0.634	3.5735
…been involved in disputes with other community members.	0.0591	0.0398	0.1375	0.312	3.2215
Over the last year have you … (1=often, 2=sometimes, 3=rarely, 4=never)					
…major disputes with a neighbor.	0.0178	0.0293	0.5446	0.634	3.6667
…major disputes with a family member.	0.0587	0.0292	0.0443	0.312	3.6673
…major disputes with a community leader.	0.0120	0.0151	0.4289	0.634	3.9203
…major disputes with police.	0.0000	0.0144	0.9980	1.000	3.9370
…been in a physical fight.	0.0395	0.0216	0.0678	0.312	3.8233
5) Risky Behavior Index	-0.0043	0.0416	0.9173	1.000	0.0000
Over the last year have your friends … (1=often, 2=sometimes, 3=rarely, 4=never)					
…gambled or bet.	0.0022	0.0348	0.9495	1.000	3.5743
…smoked cigarettes.	-0.0055	0.0153	0.7215	1.000	3.9392
…drank alcohol.	0.0176	0.0369	0.6332	1.000	3.4006
…smoked marijuana.	0.0109	0.0149	0.4617	1.000	3.9315
…used other heavy drugs.	-0.0034	0.0115	0.7635	1.000	3.9671
…sold drugs.	-0.0145	0.0123	0.2369	1.000	3.9689

Note: All indices coded so that positive="better" outcomes. Welfare today and all indices have been standardized based off of the control group. The Psychosocial Index (PSI) is comprised of the normalized sum of Welfare today, Locus of Control Index, Self-Esteem Index, Aggression Index, and Risky Behavior Index. Q-values are calculated following Anderson (2017) to adust for multiple hypothesis testing. The five components of PSI are adjusted for testing five hypotheses simultaneously. Q-values for the underlying questions, listed below each sub-index, are adjusted by group of underlying questions, such that the locus of control index uses the P-values of the eight questions in column (3) to calculate the reported Q-values in column (4). All regressions also include controls for age and age-squared, as well as dummies for gender, educational attainment, and community fixed effects.

Table A.9: Life Events and SFC Treatment Status

	Treatment	Control	P-Value of Diff
Life Event Index (Standardized)	0.1440	0.1654	0.6219
	(1.0019)	(0.9752)	

Have any of these things been a worry for you or anyone else living in this house during the last year? (1=yes, 0=no)

	Treatment	Control	P-Value of Diff
Serious Illness *	0.5173	0.5260	0.6927
	(0.4999)	(0.4996)	
Serious Accident *	0.2726	0.2726	1.0000
	(0.4455)	(0.4455)	
Death of Friend/Family Member *	0.5857	0.6252	0.0657
	(0.4928)	(0.4843)	
Divorce or Separation *	0.2380	0.2459	0.6728
	(0.4260)	(0.4308)	
Lost Job	0.4629	0.4682	0.8077
	(0.4989)	(0.4992)	
Not Able to Get a Job	0.7454	0.7620	0.3798
	(0.4358)	(0.4260)	
Alcohol Related Problems	0.2551	0.2486	0.7326
	(0.4361)	(0.4324)	
Drug Related Problems	0.1668	0.1551	0.4678
	(0.3730)	(0.3622)	
Witness Violence	0.3501	0.3516	0.9414
	(0.4772)	(0.4777)	
Abuse or Violent Crime	0.2604	0.2842	0.2229
	(0.4390)	(0.4513)	
Trouble with Police	0.2054	0.2102	0.7868
	(0.4042)	(0.4077)	
Gambling Problem	0.1863	0.1956	0.5912
	(0.3895)	(0.3968)	
Family Member Sent to Jail *	0.1950	0.1994	0.7995
	(0.3964)	(0.3998)	
Overcrowding at Home *	0.2951	0.2938	0.9503
	(0.4563)	(0.4557)	
Discrimination/Racism *	0.2539	0.2534	0.9796
	(0.4354)	(0.4352)	
Vandalism *	0.2217	0.2017	0.2670
	(0.4156)	(0.4015)	

Notes: Life events with an asterisk (*) are included in the aggregate life event index. Standard errors in parentheses.

Table A.10: Baseline PSI Leave Out Mean Heterogeneity on Low Predicted LFI

	(1) Ever Attend	(2) Total Attendance	(3) PSI	(4) LFI
SFC	0.730***	10.735***	0.042	0.015
	(0.019)	(0.306)	(0.058)	(0.051)
SFC x Low Pred LFI	0.020	0.307	-0.121	0.210***
	(0.024)	(0.353)	(0.080)	(0.072)
Group Leave Out Mean	-0.149	-1.715	-0.085	0.419***
	(0.092)	(1.249)	(0.133)	(0.153)
Group Leave Out mean x Low Pred LFI	-0.321*	-4.823*	-0.010	-0.396*
	(0.166)	(2.412)	(0.490)	(0.221)
Low Pred LFI	0.014	0.293	0.073	-0.172***
	(0.021)	(0.320)	(0.084)	(0.056)
LFI - Lagged	-0.007	-0.077	0.024	0.091***
	(0.008)	(0.115)	(0.026)	(0.024)
PSI - Lagged	-0.006	-0.102	0.174***	0.041*
	(0.008)	(0.128)	(0.024)	(0.021)
N	2007	2007	2007	1947
Mean Y	0.3687	5.4210	-0.0119	0.0539
Control Mean	0.0000	0.0000	-0.0000	-0.0033
R2	0.5964	0.5879	0.0676	0.0816

Notes: Group leave out mean is calculated as the average baseline PSI level of group members, excluding oneself. It is calculated to be centered on zero for each community, thus showing standardized deviations from community averages. Low Pred LFI indicates low predicted LFI outcomes. $* p < 0.10$, $** p < 0.05$, $*** p < 0.01$

Table A.11: Shue (2013) Peer Effect Analysis Estimation

	Psychosocial Index				Labor Force Index			
	(1) Raw Diff	(2) Resids	(3) Ancova	(4) Ancova	(5) Raw Diff	(6) Resids	(7) Ancova	(8) Ancova
Same Group	0.007	0.010	0.010	0.010	0.047***	0.053***	0.053***	0.053***
	(0.007)	(0.008)	(0.007)	(0.026)	(0.007)	(0.007)	(0.007)	(0.020)
Constant	1.103***	1.086***	1.067***	1.067***	1.142***	1.117***	1.108***	1.108***
	(0.002)	(0.002)	(0.002)	(0.023)	(0.002)	(0.002)	(0.002)	(0.019)
N	281306	275235	275235	550470	264350	258459	245514	491028
Mean Y	1.1031	1.0871	1.0677	1.0677	1.1449	1.1205	1.1118	1.1118
Distance Ratio	-0.0059	-0.0090	-0.0093	-0.0093	-0.0414	-0.0475	-0.0480	-0.0480
SE Clustering	Obs	Obs	Obs	Two-Way	Obs	Obs	Obs	Two-Way
R2	0.0000	0.0000	0.0000	0.0000	0.0002	0.0002	0.0002	0.0002

Notes: "Two-Way" indicates implementation of two-way clustering with dyadic data following Fafchamps and Gubert (2007). "Raw diff" indicates an outcome variable that is the raw difference in outcomes for each dyad. "Resids" is the difference in residualized outcomes after removing variation explained by the basic covariates used throughout the analysis (gender, community ID, education, age, and age-squared). "Ancova" indicates that the lagged outcome variable (PSI or LFI) is included in the residualization process before calculating the difference in the remaining variation.

Appendix B: Sports for Development Linked Organizations and Expenditures

In order to identify active organizations working on SFD related programs we began with the available lists of SFD programs located at `https://www.sportanddev.org` and `http://www.streetfootballworld.org`. These organizations list 1049 unique organizations running SFD programs. However, these lists may be incomplete and include a wide range of actors who operate at widely ranging scales and capacities. With such a broad umbrella, and without a centralized clearinghouse of SFD focused programs and organizations, a reliable and comprehensive estimate of all actors in this space is not possible. This makes estimating the magnitude of global expenditures on these programs extremely difficult. Organizations that can be identified do not always make financial records publicly available. Documents that are available are frequently incomplete or do not clearly distinguish allocations to sports activities from other programmatic expenditures. Moreover, SFD activities are often embedded within larger programs or funded through corporate social responsibility, making accounting for SFD expenses opaque.

In 2003, the International Platform for Sport and Development was formed as a way to link sport and development organizations. The IPSD has eight governmental and NGO leaders on its steering committee: The Commonwealth Secretariat, the Norwegian Olympic Committee, the Union of European Football Associations, the Australian Department of Foreign Affairs and Trade, Laureus Sport for Good Foundation, Reach Out to Asia, and the German Federal Ministry for Economic Cooperation and Development. We were able to connect these eight organizations alone to over $105 million in project budgets for SFD-related activities. We also identified a number of other actors with publicly available disclosures of their SFD expenditures including FIFA ($7 million), streetfootballworldwide ($2 million), the Inter-American Development Bank ($4.5 million), Clinton Foundation ($8 million), Right to Play ($48 million), and a World Bank project with the

13

Russian Federation valued at $150 million. Although specific funding numbers could not be tracked down, our review also found a number of other programs and organizations with SFD funding or programming including UNICEF, UNESCO, America SCORES, the Foundation for Global Sports Development, International Olympic Committee, Charity Ball, the Clinton Foundation, The Commonwealth, Right to Play, UEFA, and USAID.

Appendix C: Recruitment and Randomization Protocols

In order to mobilize youth for the program, Mercy Corps and its implementing partners raised awareness of the registration date in each community. In the days leading up to the event Mercy Corps publicized the event by circulating fliers, posting large banners throughout the community, and informing local authorities. On the day of the registration itself, Mercy Corps used a truck with a large amplification system to drive through the streets of the community, broadcasting information about the recruitment and encouraging youth to go to the registration location and sign up. This broadcasting approach began early in the morning of the registration day and extended until the targeted number of registrants had been reached, often late in the afternoon or early evening.

At each registration center, typically a local school or community center, interested youths queued in line, waiting for admission in the order in which they arrived. They were then allowed into the registration room in sets of approximately 30 and explained the details of the potential lottery outcomes. After completing registration, individuals then chose their assignment ticket which included whether or not they were in a SFC group and, if so, which team they were assigned to. The tickets were chosen from a covered bucket so that they could not influence their selected outcome. The result of their draw was then recorded by research team staff before they were given additional information about their selected program and group.

Appendix D: Sport for Change Program Curriculum

Mercy Corps Liberia designed the SFC curriculum based on curricula employed by Mercy Corps' international Moving Forward program. The approach covered five life skills competencies: (1) constructive communication, (2) self-esteem, (3) resilience and problem solving, (4) teamwork and trust building, and (5) strategy making and planning. Twelve of the sixteen group meetings were designed to cover at least one of the five life skills competencies. Table 4 shows the distribution of life skill competencies coved in each session. The remaining four meetings provided the group with an opportunity to prepare for an end-of-the-program graduation ceremony. Coaches were instructed to follow the prescribed structure of session meetings. Mercy Corps representatives audited compliance with the session guidelines and provided assistance where necessary.

The general session structure included an introduction period, aerobic warm ups, 2–4 group activities centered on the life skill competencies highlighted for the session, a team sports activity, and a concluding period. During each stage in the session, coaches led the initial discussion or demonstrated an activity then encouraged discussion among the group participants. Figure A.1 shows a prototypical schedule for an individual SFC group session.

The topics covered in each of the sessions:

Session 1: Introduction

Topics covered:

- Getting to know each other

- Understanding the goal of Sports for Change

- Making a Club name

- Making goals for the Club

- Developing a code of conduct

Session 2: Me and Others Part 1

Topics covered:

- Finding meaning in others

- Understanding that others are different

- Respect

Session 3: Understanding Emotions

Topics covered:

- Understanding and recognizing emotions in others

- Communicating emotions ourselves

- Dealing with heightened emotions like anxiety, stress, and elation

Session 4: Communications

Topics covered:

- Importance of communication

- Methods of verbal communication

- Methods of non-verbal communication

Session 5: Relationships

Topics covered:

- Peer pressure

- Trust

- Relations with adults

Session 6: Me and Others Part 2

Topics covered:

- How to engage effectively with others

Session 7: Believing in ME

Topics covered:

- Self-confidence

- The importance of others in being us

Session 8: Conflict and Violence

Topics covered:

- Why do conflicts happen?

- How do we avoid conflict?

- Fair play

- How do we deal with conflicts when they arise?

Session 9: Me and Others Part 3

Topics covered:

- How to plan to work well with others

- Effective group communication and strategizing

Session 10: Motivation

Topics covered:

- Concentration

- Working hard

- Not getting down when times are tough

Session 11: Dealing with Problems

Topics covered:

- Resilience

- Problem solving

Session 12: Understanding Emotions

Topics covered:

- Planning

- Thinking before doing

- Cooperative strategies

Session 13: Applying SFC Skills in My Life

Topics covered:

- Applying Sports for Change Skills in our lives

- Discussing end of project and Graduation Event

- Training in Planning

Session 14: Planning Community Graduation Event

Topics covered:

- Planning for the Graduation Event

- Establishing a planning committee

Session 15: Sports for Change Skills Refresher

Topics covered:

- Review of all 5 Sport for Change skills

- Update of the Community Graduation Event planning

Session 15: Review and Closing

Topics covered:

- Review of Graduation Event

- Evaluation of Sport of Change program

- Answering any further questions

Appendix E: Psychosocial Measures and Sub-Indices

We measure subjective welfare using a modified version of the Cantril Scale (Cantril 1966). Respondents were shown a six-step ladder and told that the top of the ladder represents the wealthiest households in the community while the bottom represents the poorest. They were then asked to indicate which step of the ladder they saw themselves at with one being the lowest and six being the highest response values. Deaton (2008) has shown that the Cantril Scale of subjective welfare correlates with income levels. An average of 2.3 in the sample suggests that, on average, respondents felt that their overall welfare was just below average.

Along with this measure of subjective welfare, locus of control, aggression, risky behavior, and self-esteem comprise the four other psychosocial well-being measures we use in our analysis.[27] These outcomes were drawn from internationally accepted psychosocial indicators. Each of these indices is the composite scoring of a series of questions on

---

[27]Although we collected data at baseline, the depression anxiety and stress (DASS) index was excluded from the endline due to challenges in implementing the module over the phone.

respondent attitudes and behaviors linked to their feelings of control for their life outcomes (locus of control), self-esteem, risky behaviors, and aggression.

Locus of control, developed by (Rotter 1966) is a concept of internal versus external control. High values of the Locus of Control Index signify that an individual believes that his or her actions determine the outcomes, whether positive or negative, in their life. Lower values indicate external locus of control, meaning that an individual believes that his or her behavior has little influence over the outcomes in life. The index involves a series of eight statements that the respondent is asked to identify with, e.g. "If you try hard you can make your life better." For each statement, the respondent is asked to state if they Agree, Somewhat Agree, Somewhat Disagree, or Disagree. To create the index, responses are sorted and scored from 1 to 4, with external control responses equaling 1 and internal control responses equaling 4. The values of the eight responses are equally weighted and summed to create an index that ranges from 8 to 32 with higher values suggesting greater internal control of personal outcomes.

The Rosenberg Self-Esteem Scale (Rosenberg (1965)) measures positive and negative feelings about oneself. We implemented the scale using 8 statements which the respondent is asked to use a Likert Scale for Agree-Disagree. Negative statements were reverse coded and total score calculated for each respondent by summing the responses to the 8 statements. As with the Locus of Control Index, the Self-Esteem Scale ranges from 8 to 32 with higher values indicating higher levels of self-esteem.

In order to identify tendencies toward aggressive behavior, we presented seven distinct scenarios to respondents. Scenarios were framed over a 12 month recall period, e.g. "In the past 12 months have you had major disputes with a neighbor often, sometimes, rarely or never?" Responses were coded so that "Never" equals zero and "often" equals four. Scores for the seven questions were summed in order to create an Aggression Index ranging from 0 to 21. At baseline, the vast majority of respondents did not report aggressive behavior, with a mean index score of 2.55 and median of 2.0. However, 10% of

respondents have scores of 6 or more, suggesting that a small segment of individuals are involved in altercations within their community.

We also asked a series of six questions about risky behaviors (gambling, cigarette smoking, alcohol use, marijuana use, hard drug use, and drug selling). Each question was coded from Never (0) to Often (4) and all six are equally weighted and summed to create an index ranging from 0 to 18.[28] In later regression-based analysis, we invert the indices on aggression and risky behaviors in order to be consistent with "better" behaviors captured by more positive values. We additionally standardize all psychosocial sub-indices as well as the aggregate psychosocial index, PSI.

Appendix F: Summary of Qualitative Research Findings

From March to May of 2014, a team of young adults in four communities were trained to complete ethnographic journals. Ethnographers were instructed to record information about events or activities that concern the youths in their respective communities. Our aim was to understand what youth were talking about in each of the week the diary was written and how events within the community were reflected through conversations.

The journals were written on a weekly basis for a period of nine weeks. The ethnographers were instructed to record events, disputes, conversations, and attitudes in the community. Ethnographers were provided a series of prompts in order to help motivate journal writing. These prompts included:

- What are the youth in your community talking about this week? (e.g. gossip, events inside and outside the community)

- How are youth in your community feeling this week? (e.g. are people worried about anything, excited about anything, angry about anything?)

---

[28]Our measures of aggression and risky behavior follow the survey modules used in Blattman, Jamison, and Sheridan (2017) who adapted the aggression module from Raine et al. (2006).

- What things are youth in your community doing this week? (e.g. programs within the community, work, leisure activities, gatherings)

- Has anything happened this week that affects youth in your community?

The purpose of the ethnographic journals was to provide qualitative context to the day-to-day lives of young adults in urban Monrovia. Each ethnographer wrote a three-page weekly journal and submitted them to a field research manager. A total of sixty-one journal entries were collected by the end of the nine week period. These diaries were scanned and transcribed by experienced data entry personnel.

The ethnographic journals allow us to provide more context- and time-specific description of the lives of youth in Monrovia. Moreover, the structured survey interviews from the second round of the panel survey provide insight into coping strategies at the household level, while the journals provide insight into community and collective strategies for coping with adverse events.

Our analysis of the ethnographic journal involved a series of steps to prepare and analyze the data. First, we identified a list of top-level codes that are specific to events and living conditions that we expect to see in the journals. Our initial codes were derived from the quantitative analysis, from which we could identify life events commonly reported by youth through survey responses. Survey analysis pointed to life events, living conditions, and shocks as key themes that would, in theory, affect the social environment of youth. We employed a two-level method for coding of the qualitative journals. In the first stage, we created three themes to explore in the journals: (1) living conditions, (2) community events, and (3) unexpected events (shocks). In the second stage, within each of these themes, we derived a list of codes for document analysis.

Our analysis of the ethnographic journals suggests that youth are commonly confronted by death, fights and/or violence, drug-related problems, theft, and/or crime. Figure A.2 displays the distribution of coded events in the ethnographic journals. We note that higher reported adversity in a community does not necessarily indicate a incidence

of these events in a community, but instead higher occurrence of codes suggests that a particular event is more commonly discussed and therefore included in the journals.

Death

Emotional problems are commonly mentioned in the ethnographic journals. In all four communities we analyzed, such problems tend to be linked to death. Our survey data, both before and after the Ebola crisis, point to a high occurrence of death. During the 2014 survey round 73% of young adults mentioned that a family member or close friend died in the preceding 12 months. This figure was lower, but still 61% when asked in 2015. Interestingly, the prevalence of death appears lower in the year when Ebola occurred, which suggests that life in Monrovia is filled with great risk to life, even without the presence of Ebola.

Death in the community generates fear among youth. Youth mourn lost friends, family, and community members, as described in this passage:

> *Youth are not feeling good this week because of the death of the out–going president for the youths child. Youth are really affected because they are somehow sad and not really going to do their normal business because of the death of their out-going president child. They all are mourning his lost.* [Banjor, Week 3, J1]

Another example of the ever-present concern about death points to the social grieving process that youth go through.

> *They later found the boy lying dead in the water. This situation has sadden the youth and the entire community that has also effected them: stop their social activities were in on a Sunday they go on the beach to mere make and have fun with some of their friend.* [Banjor, Week 1, J2]

Infrastructure

The second most frequently occurring code is community infrastructure, specifically roads. Young adults mention their embarrassment over the bad road conditions in their community. They appeal to the government to improve the bad roads and build new ones.

> *Youth, are talking about these constructions [roads] and some are even calling on the district radio and thanking the Lawmaker at the same time calling on him to continue his good work. Mean while, others are condemning the protect saying that they are substandard and after wards feeling disappointed in their lawmaker.* [Logan Town, Week1, J7]

In one journal, youth discuss the roll that bad roads play in delaying the delivery of health services to pregnant women, which leads to the tragedy of the death of their babies. The youths would feel very excited when new roads are built because they believe convenient roads would improve their community's quality of living.

> *These situations have lead many pregnant women to lose their babies due to the location of the clinic . . . This also cause lots of side effect to pregnant women while travelling from their various houses to the clinic using the motor cycle along the bad road.* [Chicken Soup Factory, Week 3, J3]

Another sign of overtaxed infrastructure is the 29% of survey respondents who state that they have experienced overcrowding at home. Dwellings tend to be very small. Our survey data suggests that the typical household that Monrovia household lives in has 3.1 rooms, thus the average room is occupied by 2 or more people.

24

Fights or violence

Fights and violence are also a common problem in Monrovia. Besides the common fights, such as among motorcyclist, fights and violence in Monrovia also include altercations among young women and men, and between parents and children. For the latter situation, the youths are very concerned about the children harmed by a parent:

> *Youth are talking about the woman who beat her daughter and broke her arm. According to my fellow youths in a small inter-view with them, this woman is very heartless and doesn't have feelings for children. She even needs to be punished for this by youths.* [Logan Town, Week 4, J8]

Approximately one-quarter of youth mention that they or someone in their household has experienced abuse or violent crime the past 12 months. We observed a slight increase in the prevalence of violent crime between round 1 and 2 of the panel survey.

Drug and alcohol related problems

Drug use is a serious problem among youth in Monrovia, leading to misbehavior or even fights. The survey data suggests that drug use is a problem among approximately 16% of households where young adults live. Alcohol is more commonly cited, 27%. However, the importance of drug-related problems mentioned in the ethnographic journals is likely due to the heightened severity of problems with in the community, in particular, among youth.

Young adults look to police to deal with drug-related problems in neighborhoods. The journals mention that youth will do everything possible to encourage the government to prevent the use of alcohol and drugs.

Taking in drugs, smoking grass has become a very bad habit of some youth the dry in the community this week. This has brought to the attention of several youth in the community to hold a series of discussion concerning a plan to find a solution to this situation. [Dry Rice Market, Week 4, J5]

Theft or crime

Vandalism is cited by 15% of respondents in round 1 of the panel survey and 21% in round 2. Theft and crime underscore the prevalence of violence and insecurity within urban neighborhoods of Monrovia. For instance, the example below describes that a person was murdered because of theft. Youth employed violence to punish the thief. The response points to the use of vigilante justice and indicates the need of a proper and legal way to deal with the theft.

In the community a boy was murder for theft. The alleged victim was seen laying in the drainage in the community ... [Banjor, Week 2, J2]

Coping strategies

Young adults utilize several coping strategies in response to the challenges that they face in their daily lives. Specifically, youth talk in groups about both regular and extraordinary events. They organize into intellectual forums and advocate for change on topics and issues which they feel strongly about. Repairs to infrastructure are among the central uniting motivations for young adults. From calls for collective action in reducing garbage and improving sanitation in community to youth calling for action to improve roads, youth display their proclivity towards civil action. From our analysis of the ethnographic journals, we summarize coping strategies in two groups:

1. Local social action - youth acting through collective organization

2. Calling for government assistance - social action that youth where youth have identified a need for government involvement

We explore the two coping strategy methods below.

Local social action

When shocks occurred, youth would usually gather together and discuss the possible ways to solve the problem. Sometimes they would pray. In one instance a journalist mentions youth switching from football to prayer/church in response to the threat of the prayer mother.

> *This is something that has caused the youths some emotional problems. It has affected their Sunday morning football practices because most of the youths that usually practice on Sunday morning no longer go for practice, but rather go for prayer service to the prayer mother.* [Banjor, Week 1, J1]

In resolving intra-households disputes, youth discuss their strategies to reach resolution with parents and guardians.

> *Finally, more than half of the youths think that parents should visit these entertainment centers at night and see for themselves. They say, with this, they and their parents can brain storm and remedy the problem. [Dry Rice Market, Week 1, J5]*

However, when families are unable to resolve their disputes, youth may join efforts to support their peers, as suggested by the following quote, in which the journalist discusses an effort to collect money from community members in order to resolve a family dispute over land.

Youths are talking about the families that are fighting among themselves for a piece of land. This discussion to place in front of the video club at the gas-station. To stop fight, young people of the community are collecting money to bring surveyor and survey the land to give every one the share. [Banjor, Week 2, J1]

In other circumstances, youth resort to more aggressive social action. A particularly illuminating example of the passion that young adults demonstrate comes from the following journal sample, where the journalist mentions the surging passion of youth in a call for justice. The journalist mention the organization of an "angry youthful crow" to symbolize the collective action of youth within the community.

A respectable Christian mother had just confessed her long time link with the devil for demonic & spiritual powers. [Reverend Mother] ...a pastor and owner of the fire healing and Deliverance ministries ...is said to have made blind a young man ...when he openly try to exposed her evil secrets. [He] live very close to [Reverend Mother's] compound. It is said that [Reverend Mother] took an ordinary sand which she chanted upon and wasted it into his which got him blind on the spot after she suspected him as being the criminal who entered her compound to steal her cars tyres. Many youth are angry about this and are at the same time asking the community leaders to vanish the demonic mother from their community. The family members of has taken the Rev Mother to the Magistrate Court for trial. As a living witness, I swathe Mother lively testifying at court which was on May 3, 2014 @ 10:00 am surrounded by angry youthful crow, since [the man] is a young man to conclude[29]*, they are feeling bad and disappointed.* [Logan Town, Week 1, J6]

Young adults are much more active in taking actions than we initially expected. They are willing to take the responsibilities to serve their communities. In the following quote, youth form a community watch in response to theft.

---

[29]In this context the journalist uses the term "conclude" to mean the pre-mature death of the young man.

Some youth of the community begin to investigate who must have done such an ugly art (theft). ...The youth of community have decided to put in place a Community Watch team in order to fight crime especially at night. [Chicken Soup Factory, Week 1, J3]

Where local governance and adult leadership is insufficient, youth sometimes take it upon themselves to organize collective action to repair infrastructure.

Due to the failure of the community leaders, the youth themselves took up the responsibility to re-condition the road using local tools such as hoe, shovel, dicker, cutlass and other related tools. [Chicken Soup Factory, Week 3, J3]

Calling for government assistance

We observe that Monrovian youth even repair roads by themselves using local tools when noticing the government inaction. Youth appeal to government to do something about their situation, especially for the public infrastructure, such as roads and public toilets. They employ very smart and strategic way to encourage the government to improve the infrastructure. For instance, they thank a lawmaker on the district radio and call on him to continue the contribution. They also urge the government to empower the youth of the community, showing their strong willingness to take part in serving their communities. For drug and crime relate problems, they will find the police to help.

They told me that they are feeling very disappointed in community dwellers who built houses without better latrine. They also stressed the need for the community chairman to ensure that every community permanent resident to build a latrine for his or her house, if all effort to have the Representative build the latrines should fail. [Dry Rice Market, Week 1, J5]

Direct communication with government and non-governmental representatives is a common strategy of adolescents.

> *The youth is kindly asking the community leaders to engage Government and NGOs*
> *to build another clinic to a strategic location to accommodate the entire community.*
> [Chicken Soup Factory, Week 3, J3]

Youth also coordinate action to involve and inform the police, especially in the case of drug abuse.

> *Many young boys and girls of the community are involved in the abuse of drugs*
> *there by causing problems in the community. The young people are now giving this*
> *information to the national police of liberia to help in the fight of this ugly behavior.*
> [Chicken Soup Factory, Week 4, J3]

Session 8: Conflict and Violence

Topics to be covered:
- Why do conflicts happen?
- How do we avoid conflict?
- Fair play
- How do we deal with conflicts when they arise?

Equipment
- Ropes
- Footballs – 6
- Bibs
- Cones
- Other sporting equipment as required by the sport

Session Outline

0-5mins	Energizer	Opening Chant Travelling to new places
5-20mins	Introduction	Review of last session Feedback from last session Introduce topic for the day Outline activities
20-35mins	Warm up	Stretches and exercise
35mins-1hr	Activity 1	Knee Fight
1hr-1hr20mins	Activity 2	Find Similarities and Move On
1hrs20mins-1hr45mins	Activity 3	Line Push and Pull
1hrs45mins-2hrs	Break	Including registration
2hrs-2hrs40mins	Sport	Football (2 varieties to promote conflict resolution!)
2hrs40mins-3hrs	Conclusion	Review main topics discussed Discussion Plan for next week Closing Chant

Figure A.2: Distribution of event, living condition, and shocks codes in journals