Network Partitioning and Social Exclusion under Different Selection Regimes

Clara Delavallade, IFPRI

Alan Griffith, University of Michigan

Rebecca Thornton, University of Illinois at Urbana-Champaign

This Draft: March 1, 2016

Abstract

While most social programs are based on some form of exclusion of sub-populations, we know little about how being excluded, and the selection process, affect social inclusion. This paper compares peer effects of an after-school program, under three different (randomly assigned) network-formation regimes: endogenously formed, popularity vote, and randomly assigned. We find substantial evidence of homophily within endogenously-formed and elected networks. When participation was randomly assigned, we find segregation of friendships due to the program. We do not find this among elected networks, mainly because they were already highly partitioned. Lastly, we find that social exclusion – not being elected in a school with popular voting – reduced education aspirations and self-confidence.

1. Introduction

The literature on social networks within Economics has primarily focused on either quantifying the causal effects of peers on outcomes (Altermatt and Pomerantz, 2003; Bearman and Moody, 2004; Fowler and Christakis, 2008; Burgess and Umana-Aponte, 2011) or understanding how networks are formed (Burgess et al. 2011). However, little attention has been paid to the interaction of these two processes. In this paper, we examine how the network formation process itself affects how peers impact others. To this end, we compare peer effects under three different (randomly assigned) network-formation regimes: endogenously formed, popularity vote, and randomly assigned. We use two rounds of network data from 30 schools in rural India to identify changes in pairwise links between students over the course of one academic year. We also utilize two levels of randomization to separately identify the causal effect of peers, of network formation regime, and their interactions.

The paper uses data collected from students in grades 6-8 in 30 schools in rural Rajasthan, centered around a girls' after-school program implemented by a local charity organization. Prior to the study, baseline surveys and network data were collected. The thirty schools were randomly assigned to three treatment arms in which girls were either voted into the program by popular election, randomly assigned to the program, or did not receive the program at all. To identify counterfactual elected girls, we conducted popular elections in each of the 30 schools prior to program randomization. At the end of the school year, we conducted another round of questionnaires and network surveys.

We first examine networks at baseline under the three formation regimes: endogenously formed networks, popularity vote, and randomly assigned networks. Consistent with a large literature on sorting (Kandel 1978; Hamm 2000; French et al. 2003; Burgess et al. 2011), we find substantial evidence of homophily within endogenously formed networks. Two girls are more likely to be friends with each other if they share characteristics in common, such as being in the same grade or the same age. Elections led to even similar and tighter networks. Those elected are significantly older and in a higher grade than those who were not elected. Further, election results show substantial evidence of endogenous sorting, as two elected girls are 24.9 percentage points more likely to be friends at baseline than two non-elected girls and 15.0 percentage points more likely to be friends than if only one is elected. In contrast, we find that random assignment was successful in creating balanced groups of selected and non-selected girls.

We then examine network formation and changes, under each selection regime, after the after-school program has run for approximately four months. We find that endogenously formed friends at baseline are substantially more likely to still be friends at the endline. In schools with participants chosen by election, there is no added effect of being elected on the likelihood of friendship at endline. However, we find some evidence of segregation between girls who were randomly selected for the program and girls who were not

selected. Two girls who were not selected are 16.7 percentage points more likely to be friends at endline. Two girls who were selected for the program are 20.8 percentage points more likely to be friends at endline. In contrast, girls of whom only one of the pair were selected to participate in the parliament program are 6.8 percentage points less likely to be friends at endline than if neither had been selected.

Lastly, we turn to measuring the causal effects of the after-school program under each selection regime, on education and career aspirations, self-confidence, and gender roles attitudes. We find that being in a school with popular voting reduced education aspirations and self-confidence overall, and that these effects are mainly driven by students who were not elected. Non-elected girls in schools with the program that had popular voting, have a self-confidence index 0.37 standard deviations lower than those in the control group, suggesting a possible discouragement effect from not being elected. We do not find this effect on girls who were randomly selected. In addition, we find that exclusion affected even those who were ineligible for the program. Boys in both treatment arms have a lower self-confidence at endline than those in the control group.

This paper makes a significant contribution to a nascent literature that accounts for network dynamics in measuring peer effects. Over the past two decades, a growing literature in economics and related fields has investigated the importance of one's peers to a large variety of economic and social outcomes (See, e.g., Miguel and Kremer 2004; Oster and Thornton 2012).[1] A severe limitation of this literature is that it almost uniformly assumes that networks are static, or at least exogenous. This assumption may be innocuous in settings where networks are indeed random (De Giorgi, Pellizzarri, and Redaelli, 2010; Sacerdote, 2001), or when interventions are unlikely to affect network structure (Ngatia, 2011). However, a large literature in sociology and related fields demonstrates that links are far from random. Importantly, social networks tend to demonstrate *homophily*, whereby individuals are more likely to be friends with individuals similar to them by race, age, gender, etc. (See, e.g., Currarini, Jackson, and Pin, 2009). Failure to account for endogeneity of networks may lead to biased estimates of peer effects.[2]

In addition, there has been very little research accounting for changing network structure. In a recent paper, Comola and Prima (2014) investigate the effect of randomized access to savings accounts, accounting for changes in network structure due to their intervention. As in our setting, they collect data on network structure pre- and post-intervention, so as to assess the effect of their intervention on the network

[1] Networks have been shown to affect technology adoption in many settings (Oster and Thornton, 2012; Conley and Udry, 2010; Bandiera and Rasul, 2006), and information diffusion through a network depends critically upon network structure (Banerjee et al., 2012). Who one knows is also crucially important for job referrals (See, e.g., Beaman and Magruder, 2012).

[2] Another approach to identifying peer effects with endogenous networks is to measure the effects of friendship exposure to a randomized treatment (Oster and Thornton, 2012). In the United States, the availability of AddHealth network data has allowed researchers to examine peer effects within non-random network structures.

itself. Similarly, Vasilaky and Leonard (2014) investigate an intervention directly targeting social ties among female cotton growers, demonstrating that altering social networks may be a powerful channel by which to increase agricultural productivity. To our knowledge, these are the only studies that leverages randomized treatment to measure impacts on the network itself. Failure to investigate interventions' effects on networks may lead researchers to neglect an important channel whereby outcomes are determined.

Lastly, while most social programs are based on some form of exclusion of sub-populations, we know little about how the selection process affects outcomes. This paper is the first to provide rigorous evidence that the selection process matters for network formation and outcomes of a girls' empowerment program.

The paper proceeds as follows: In Section 1, we present the program background. The experimental design and data are described in Section 2. Newtwork results are shown and discussed in Section 3, program effects on self-confidence and aspirations in Section 4. We conclude in Section 5.

1.1 Background: Programs for Empowering Girls

Socio-emotional factors play a key role in explaining gender disparities in educational achievement and labor market success. Discriminatory social norms develop low levels of self-efficacy, confidence, and well-being among girls (Dercon and Singh, 2011). Limited belief in one's own ability and self-efficacy translates into low aspirations and educational goals (Bandura et al., 2001; Cohen et al., 2009) among girls, restricting their acquisition of the cognitive and non-cognitive skills necessary to enter and succeed in the labor market (Heckman and Rubinstein, 2001).

To address many of these issues, there has been increased attention on providing girls opportunities to increase aspirations, improve self-esteem and agency, and provide a supportive and safe atmosphere to produce better long-term outcomes. Many of these programs have proven successful, such as girl-friendly schools (Kazianga et al., 2012), female role models (Nguyen, 2008; Beaman et al., 2009 and 2012), or negotiation training. What is less well known is how these type of programs affect social networks. Moreover, little is known about how the composition of the group, or the selection mechanism for participation, affects participating and non-participating individuals and peers.[3]

In this paper we study an after-school girls' parliament program in rural India, that was designed and implemented by a nongovernmental organization, Educate Girls. The program targets adolescent school girls in grades 6-8 and meet several Saturdays a month to build confidence, leadership and self-esteem. Girls who participate in the parliament undergo a life skills training based on the WHO recommendations:

[3] For example, prior research has found that a merit-based scholarship program focusing only on girls also provides persistent educational benefits to boys (Kremer, Miguel and Thornton, 2004). However, boys or girls excluded from the program may also become demotivated or disengage in school.

problem solving; critical thinking; decision making; communication; self-awareness; creative thinking; interpersonal relationships; coping with stress; coping with emotions; and empathy. The program content is delivered through a series of five "games," whereby participant girls work through scenarios dealing with complex issues such as early marriage and standing up to parental authorities. Over the course of the school year, the five games are played in a well-defined sequence under the supervision and mentoring of community workers trained and monitored by Educate Girls. While the games are designed to last about one to two hours, the parliament meeting sessions usually last around 4 to 5 hours. Overall, parliament members spend an average time of 25 hours together, allowing friendships to form and develop and for the program to affect participants.

The parliament program involves a democratic popular vote, wherein 13 girls in grades six to eight are elected by their peers (including boys). The 13 positions include a president, as well as secretaries and assistant secretaries of education, sports, management, culture, health, and motivation. Each position has two nominees. Girls are either nominated or volunteer to be considered for a position in the parliament. In most cases, the election is determined by a public show of hands.

Girls participating in the parliament are encouraged to share their skills and knowledge with other girls in the school by organizing biweekly life skills-oriented games. According to Educate Girls, girls selected for the parliament through popular elections are more likely to be vocal and better socially connected, allowing for a wider diffusion of the newly acquired life skills.

2. Research Design

Baseline Data Collection and Measures

During the 2013-14 academic year, the girls' parliament program rolled out to new districts in rural Rajasthan. We selected thirty schools from two of the new administrative blocks to participate in the study. The study involved students who were in grades six, seven, and eight, in each of the study schools. In total, there were 2655 students in these grades enrolled at the beginning of the 2013-14 school year.

At the beginning of the school year and prior to program implementation, we conducted baseline surveys asking students about their background, aspirations, self-confidence, and attitudes toward gender roles. Only students who attended school on the day of the survey, 70.2 percent, have these baseline data.

In addition, on a different school day enumerators conducted a detailed network survey to collect extensive data on connections among students. In each school, boys and girls provided information on their social ties to the girls (not boys) in grades 6, 7, and 8. Time constraints prohibited collection of each student's social tie to boys. To collect the network data, each female student would stand up one at a time,

and every non-standing student would answer questions about their link with the standing girl. In total, 71.6 percent of enrolled girls and 68.6 percent of enrolled boys completed the baseline network survey.

We use the baseline survey and network data to test for balance across randomization arms, and to control for baseline measures of empowerment and network ties. To measure empowerment, we use the survey data to construct four indices for the following outcomes: educational aspirations, career aspirations, self-confidence, and gender roles. We first collapse any questions with categorical outcomes into a series of binary indicators, indicating higher aspirations, self-confidence, or views about gender. We sign these such that a positive change in the index indicates a positive change, such as desiring to get married at a later date, more self-confidence, or stating that it is okay for a wife to disagree with her husband in public. We then take the first principal component of the variables within a given category. Finally, we normalize each index such that the mean of each is zero with standard deviation of one.

The network data allows us to create links between students at each school. Because we only asked individuals to report their ties to the girls in the class, we focus our network analysis among girls. Our primary definition of a network link involves having answered "yes" to the question "is she is a friend?" We identify the following types of friendship links for individual *i* in the data:

- *OR* friends*:* either *i* or *j* identifies the other as as a friend ($L_{ijt}{}^{OR}$)
- *AND* friend: *i* identifies *j* as a friend and *j* identifies *i* as a friend ($L_{ijt}{}^{AND}$)

We use the notation L_{ijt} as an indicator for being linked at baseline (*t*=0) or endline (*t*=1), under the various link definition ($L_{ijt}{}^{OR}$, $L_{ijt}{}^{AND}$). Note also that *AND* friends are also, by definition, *OR* friends. In addition, for each girl, we summarize her total number of *AND*, and *OR*, friends.

Baseline Data

Table 1 presents the baseline characteristics of boys and girls in the sample. On average, girls are 12.3 years old, with the majority classified as scheduled tribe, caste, or other backward caste (25.5 percent scheduled caste, 12.3 percent scheduled tribe, 44.5 percent other backward caste). Among girls, 84 perent were enrolled the previous school year. Most families, 86.8 percent, own a television. A large proportion, 83.1 percent, of the girls' fathers ever attended school, with fewer, 56.2 percent, having mothers who ever attended school. On average, girls have a total of 7.8 *AND* friends (friends who both name each other), and 15.8 *OR* friends.

In comparison to girls, boys are older, in a higher grade, and much more likely to be from a scheduled caste, tribe, or other backward caste. Further, boys' parents at baseline have significantly lower rates of schooling as well as lower wealth correlates such as owning a television. These results are consistent with

a pattern of girls of lower socioeconomic status having ended schooling earlier than boys of similarly low status. That is, the data suggest that lower socioeconomic status girls drop out of school earlier as compared to boys, leading to higher wealth and parents' education on average for those who remain. We present baseline statistics disaggregated by standard in Appendix A.

Table 1 further presents baseline education, career, self-confidence, and gender roles indices for boys and girls. Boys have significantly higher education and career aspirations and expectations (0.401 standard deviations p=0.001, 0.281 standard deviations p=0.012). That these differences exist even despite the possible culling of lower socioeconomic status girls further demonstrates substantial societal barriers to women's achievement in this setting.

We see no differences between boys and girls in self-confidence (p=0.856). We find, however, that girls have a significantly more positive view towards gender roles at baseline than boys (0.248 standard deviations p=0.020) and this difference in attitude towards gender roles increases by standard, suggesting those with lower attitudes may be dropping out before reaching standard 8 (see also Appendix A).

Parliament Elections and Random Assignment

Prior to program implementation, but after the baseline survey, Educate Girls staff conducted democratic elections in each of the 30 study schools. The process followed the standard procedure for students to choose 13 girls to participate in the Bal Sabha. In most cases the election consisted of a show of hands (90 percent) and on average, the winner captured 75 percent of the vote (s.d. 0.157). Prior to the election, students were informed that there would be a lottery and that some schools would receive the parliament program with participants determined by the election, some schools would receive the program with participants determined randomly, and some schools would not receive the program. Enumerators recorded who was elected for each position and the result of the vote.

Randomization and Baseline Balance

After the conclusion of baseline data collection and elections, the 30 schools were randomly assigned to three intervention arms. Ten schools were assigned to T1 in which participants were determined by the outcome of the election. Ten schools were assigned to T2 in which participants were randomly selected. The final ten schools served as controls and did not receive the parliament program. The parliament program was then implemented over a period of approximately four months.

School-level randomization

To test for baseline balance, for each baseline demographic, empowerment, and network characteristic $BaseVar_{is}$ for girl *i* in school *s*, we estimate the following regression:

1) $BaseVar_{is} = \beta_0 + \beta_1 T1_s + \beta_2 T2_s + u_{is}$

where $T1_s$ and $T2_s$ indicates being assigned to T1 or T2 schools. We estimate these regression separately for girls and boys and cluster standard errors by school. Table 2 presents the estimated regression coefficients β_0, β_1, and β_2, and the p-values indicating the statistical significance of the joint test β_1. Note that the null that means are the same across the three treatment groups is not rejected for either boys or girls for any baseline characteristic, as indicated by the p-values presented in Columns (4) and (8). Samples are also generally balanced among our baseline measures of empowerment and networks. However, with the small number of clusters, balance may be an issue: T2 girls have significangly lower baseline career aspirations. Relatedly, T2 girls have significantly lower peer group mean career aspirations and T1 girls have marginally significantly lower baseline self confidence. Still, we control for baseline observations in robustness specifications.

Individual-level randomization in T2 schools

To validate the individual randomization of girls into the parliament program in T2 schools, we test for baseline balance across selected and non-selected girls in T2 schools. For each baseline demographic, empowerment, and network characteristic, $BaseVar_{is}$, for girl *i* in school *s*, we estimate the following linear regression, restricted to girls in T2 schools:

2) $BaseVar_{is} = \beta_0 + \beta_1 Selected_{is} + u_{is}$

where $Selected_{is}$ indicates being randomly selected for parliament participation. We cluster standard errors by school. Table 3 presents the average of each baseline outcome among those not selected, the estimated regression coefficient β_1, and the p-values indicating the statistical significance of β_1. Note that there are only ten T2 schools and thus we lack power to detect small differences, however, we fail to reject equality of means between girls selected and girls not selected within T2 schools, for most baseline measures.

Endline Surveys and Attrition

Approximately six months after the baseline surveys, the study team returned to each school to collect endline data, consisting of an endline questionnaire and network survey. 74.9 percent of enrolled students completed the endline questionnaire, while 75.0 completed the endline network survey. Among those who completed the baseline survey, 81.7 percent completed the endline questionnaire and 82.5 percent completed the endline network survey.

Across all of the surveys, we observe a total of 2773 students, of whom 2655 (95.7%) are found in administrative enrollment records.

We formally test for differential attrition by random assignment with Equations 1 and 2 in Table 4, using indicators of survey completion as dependent variables. Completion of a survey (Baseline, Baseline network, Endline or Endline network) was not significantly associated with random assignment of intervention arms by school.

There is, however, a significant relationship between survey completion and being selected for participation among T2 girls. Those who were selected were between 9.2 and 15.8 percentage points more likely to have completed survey data (Table 4, Panel B).

We further test for differential attrition across a number of baseline characteristics. Among girls and boys who were present for the baseline survey, we estimate the following regression to assess differential attrition by baseline characteristic:

3) $Complete_{is} = \gamma_0 + \gamma_1 \mathrm{T1_s} + \gamma_2 T2_s + \gamma_3 BaseVar_{is} + \gamma_4 \mathrm{T1_s} * BaseVar_{is} + \gamma_5 T2_s * BaseVar_{is} + u_{is}$

4) $Complete_{is} = \beta_0 + \beta_1 Selected_{is} + \beta_2 BaseVar_{is} + \beta_3 Selected_{is} * BaseVar_{is} + u_{is}$

The dependent variable is an indicator for individual *i* in school *s* having a completed baseline network survey, endline survey or endline network survey – each conducted after the baseline survey. We present p-values of the joint test for significance of γ_4 and γ_5 in equation 3 (Appendix B, Panel A), and analogously test for the significance of β_3 in equation 4 (Appendix B, Panel B), which would indicate differential attrition. We present estimates of equation 3 for girls and boys, while restricting estimation of equation 4 to girls in T2 schools. While a small number of p-values are below the conventional significance levels, we see no clear patterns suggesting differential attrition on any baseline characteristics or outcomes.

3. Results

Endogenous Networks

We next turn to presenting the baseline characteristics of networks under the three selection regimes: endogenously formed, elected, and randomly assigned. We first examine the characteristics of already-existing endogenously-formed networks, prior to elections and the introduction of the parliament program. Because these data were collected before selection regimes were randomly assigned, we present the data for all schools: Control, T1, and T2. We restructure the data to pair each girl with all girls in her school,

and estimate $L_{ijs0}^{(linktype)}$, the existence of a network link between individuals *i* and *j* in school *s* at baseline (*t*=0) under the each network definition ($linktype \in \{OR, AND\}$):

5) $$L_{ijs0}^{(linktype)} = \alpha_0 + \alpha_1 \left| BaseVar_{is} - BaseVar_{js} \right| + \epsilon_{ijs0}$$

$BaseVar_{is}$ and $BaseVar_{js}$ are baseline characteristics of individuals *i* and *j*. In this context, $\alpha_1 < 0$ indicates homophily in friendship networks. We continue to cluster standard errors by school.

Table 5 presents our main homophily results. Girls in the same standard are 15.4 percentage points more likely to be *AND* friends than girls one standard apart. Girl *i* is 10.2 percentage points more likely to indicate that *j* is an *OR* friend if they are in the same standard than if they are one standard apart, and similarly, 20.4 percentage points more likely than if they are two standards apart (that is, in standard 6 and 8). We also see substantial homophily on age and prior enrollment status, but less evidence of homophily on caste and family characteristics as shown. We see statistically significant evidence for homophily in self-confidence and gender roles between girls. The coefficient on *AND* in self-confidence indicates that two girls with the same degree of self-confidence are 2.3 percentage points more likely to be *AND* friends than two girls with self-confidence measures one standard deviation apart, and 4.6 percentage points more likely than two girls who are two standard deviations apart on this measure. Similarly, the coefficient on *AND* in gender roles suggests that two girls with the same gender roles index are 4.9 percentage points more likely to be friends than a pair with self-confidence measures that differ by one standard deviation.

Finally, we see some evidence of degree homophily. In the networks literature, degree refers to the number of links that a given node/individual has (See, e.g., Jackson 2008). Girls are more likely to be friends with other girls who have a similar number of *AND* friends as they do. This suggests that popular girls tend to be friends with other popular girls, while less popular girls are more likely to be friends with less popular girls.

Popular vote

Next we examine how the elections lead to selection into participation in the Bal Sabha program. Again, because elections were held in all schools before the selection regime was randomly assigned, we present results from Control, T1, and T2 schools. Table 6 compares girls who were elected to those who were not across all schools. The results show that elected girls are systematically different than non-elected ones and provide evidence of selection into participation in the Bal Sabha program in the NGO's preferred delivery model. Those elected are significantly older and in a higher grade than those who were not elected. However, those elected are no more likely to be wealthier, as proxied by TV ownership and electricity, or

to have educated parents. Further, elected girls were no more or less likely to be of Scheduled Caste or Scheduled Tribe.

Among baseline empowerment measures, elected girls are more likely to have higher educational aspirations than girls who are not elected, but not significantly different on any of the other indices. Elected girls may be more "popular" in that they have more *AND* friends (1.21, p=0.087) but fewer *OR* friends (-0.70, p=0.586) , on average than unelected girls.. Finally, we note that elected girls tend to have a much higher proportion *OR* friends (0.151, p=0.001) and their *AND* friends (0.197, p=0.000) and also elected to participate, compared to their non-elected classmates. This suggests that election resulted in clustered cliques of girls being selected.

We also can assess the extent to which girls who are elected are connected within friendship networks at baseline. We estimate the following linear regression:

$$6)\quad L_{ijs0}^{(linktype)} = \gamma_0 + \gamma_1 E_{ijs}^{OR} + \gamma_3 E_{ijs}^{AND} + \epsilon_{ijs0}$$

Table 7 presents estimates of Equation 6, showing the relationship between the elections and friendships. Since elections were held in all 30 schools unconditional on treatment status, we investigate the election results for all girls in the sample. First, two unelected girls have a 74.6 percent likelihood of being *OR* friends, and a 33.6 percent likelihood of being *AND* friends. In Column 1, the estimated coefficients, γ_1 and γ_2, suggests that a pair of girls in which one is elected and one not elected is 8.0 percentage points more likely to be friends than a pair of with two non-elected girls. In contrast, two elected girls are 15.3 percentage points (0.080 + 0.073) more likely to be friends than two non-elected girls. Column 2 of Table 7 shows that if one individual is elected, girls are 9.9 percentage points more likely to be *AND* friends. Appendix C presents these results disaggregated by treatment assignment.[5]

Randomly assigned

As discussed above and presented in Table 3, the random assignment of girls to participate in the parliament program resulted in groups that were generally balanced across baseline characteristics.

Program Effects on Networks under different Selection Regimes

[5] Comparing network links and election results across treatment schools should not result in any significant differences due to the random assignment. With our relatively small sample of schools, unelected girls are more likely to be friends in T1 and T2 as compared to the control, which we attribute to school size effects: the coefficients on T1 and T2 become insignificant when we control for school size. In particular, based on enrollment, one control school has 177 girls, while the next largest school contains only 92.

We first present intention to treat estimates of the effect of being assigned to a program school. We then disaggregate potential program effects among those elected or not, and those randomly selected, or not.

Intent to treat estimates

To measure the intent to treat effect of the program on network links, we estimate linear probability models with the following equation.

7) $L_{is1}^{(linktype)} = \delta_0 + \delta_1 T1_s + \delta_2 T2_s + \epsilon_{ijs1}$

Here, $L_{is1}^{(linktype)}$ indicates the existence of a link bewteen individuals *i* and *j* in school *s* at endline (time 1). The parameters $\delta_1$and δ_2 identify the difference in probability of a link between girls assigned to T1 and C, and T2 and C, respectively. These estimates are unconditional on election or random selection to participate in the program. In some specifications we control for baseline networks, L_{ijs0}^{OR} and L_{ijs0}^{AND} to absorb residual variation. Further, to control for possible baseline imbalance in school size, we include school size controls. We continue to cluster standard errors by school.

Table 8 presents results estimates of Equation 7. Columns (1) and (5) suggest that girls in T2 are significantly more likely to be friends in T2 schools, but these effects become insignificant when adding the full set of controls. Statistical power is definitly a concern for these estimates, although in all of the specifications predicting *AND* network links, we can reject the hypothesis that the effects in T1 are the same as the effects in T2 (Columns 5-8).

Heterogeneous Treatment Effects

While we expect the program to affect networks, the fact that it does not affect the average likelilhood of forming a friendship link among all girls in T1 and T2 is not surprising. Rather than leading girls to be more likely to be friends with all other girls, if the program causes *substitution* of friendships, the intention to treat estimates will obscure this effect. Accordingly, we expect the program's effects on social networks to depend upon whether individuals are selected (either by election or randomly) to participate in the program or not. Therefore, we estimate the following to disaggregate the program effects:

8) $L_{ijs1}^{(linktype)} = \delta_0 + \delta_1 E_{ijs}^{OR} + \delta_2 E_{ijs}^{AND} + \delta_3 T1_s + \delta_4 T1_s * E_{ijs}^{OR} + \delta_5 T1_s * E_{ijs}^{AND} + \delta_6 T1_s + \delta_7 T2_s * Selected_{ijs}^{OR} + \delta_8 T2_s * Selected_{ijs}^{AND} + \epsilon_{ijs1}$

Here $Selected_{ijs}^{OR}$ indicates that individual i in a T2 school was randomly selected to participate and $Selected_{ijs}^{AND}$ is an indicator that both girls were selected. In some specifications we include baseline network controls and school size controls. We cluster standard errors by school.

The following coefficients are of primary interest: among girls in control schools, the coefficients δ_1 and δ_2 indicate the additional likelihood of a network link if either both or one girl of each pair is elected. The results in Table 9 are consistent with the baseline network findings in Table 7 indicating that elected girls are more likely to have friendship links.

The parameter δ_3 identifies the effect of being in a T1 school – where girls who were elected participated in the program – on the probability that two non-elected girls will be linked at endline. The parameters $\delta_3 + \delta_4$ estimate the effect of being in a T1 school on having a friendship link when only one of the girls is elected. Lastly, $\delta_3 + \delta_4 + \delta_5$ estimates the effect of being in a T1 school the program on link probability if both are elected. Recall that in T1, all of the girls who are elected participate in the parliament program. We see little evidence that being in a T1 school of changing networks. The estimated coefficients δ_3 are small and statisically insignificant. Similarly, $\delta_3 + \delta_4$ and $\delta_3 + \delta_4 + \delta_5$ are small and statistically insignificant; p-values are presented at the bottom of Table 9.

In T2 schools, however, we find some evidence consistent with segregation between girls who were selected for the program and girls who were not. Parameters δ_6 through δ_8 have similar interpretations for girls selected and not selected in T2 schools. The parameter δ_8 identifies the effect of being in a T2 school – where girls were randomly selected to participated in the program – on the probability that two non-elected girls will be linked at endline. The parameters $\delta_8 + \delta_9$ and $\delta_8 + \delta_{10}$ estimate the effect of being in a T2 school on having a friendship link when only one of the girls is elected. Lastly, $\delta_8 + \delta_9 + \delta_{10} + \delta_{11}$ estimates the effect of being in a T2 school the program on link probability if both are selected.

We first focus on Column 1. Here, a pair of girls not selected for the program is 19.0percentage points more likely to be *OR* friends at endline in T2 than girls in Control schools. This effect remains positive and significant as additional controls are added in Columns 2-4. The results have the same sign for *AND* friendships in Columns 5-8 but estimates lose significance as additional controls are added.

Among girls of whom only one of the pair were selected to participate in the parliament program, being in a T2 school reduces the likelihood of being *OR* friends at endline by 6.5 percentage points (Column 1), as compared to the case when neither is selected. This effect is robust to additional controls in Columns 2-4. Further, we note that the point estimates for *AND* friendships in Columns 5-8 are quite similar but statistically insignificant.

Lastly, if two girls are both randomly selected to participate in the program, they are significantly more likely to be friends at endline. This positive and significant result holds when adding addtional controls and with different link type definitions.

These results suggest that randomly selecting girls for the parliament program leads to segregation between those selected and those not. That is, girls who are randomly selected for the program are more likely to be friends with other participants. Similarly, girls who are not selected are more likely to be friends with other girls who are not selected.

We do not see this pattern of segregation among girls in T1. We attribute this to the fact that at the baseline, elected girls are already more likely to be friends with each other, as shown in Table 7. Because of this, there was much less space for new link formation among participants in T1 schools. Similarly, non-elected girls were more likely to be friends with each other in T1, leading to less room for new link formation among these girls in T1 schools. That is, the endogenously-formed networks that existed at baseline were largely unaffected by the treatment in T1, in which girls particpated or were excluded from participation by a pattern consistent with these preexisting networks.

Link Formation vs. Retention

The results in Table 9 provide substantial support for the hypothesis that selection and non-selection in T2 schools serves to partition friendship groups. To further investigate this, we take a further look at the effects in T2 schools broken down by baseline friendship status. To simplify the analysis, we restrict attention to symmetric link definitions AND and OR. First, we define three baseline sitautions for individuals *i* and *j* at baseline: (1) not OR friends, (2) OR friends but not AND friends, and (3) AND friends. We then estimate Equation 9 separately for pairs in each situation.

9) $$L_{is1}^{(linktype)} = \delta_0 + \delta_1 T2_s + \delta_2 T2_s * Selected_{ijs}^{OR} + \delta_3 T2_s * Selected_{ijs}^{AND} + \epsilon_{ijs1}$$

In this specification, δ_1 identifies change in probability of a link existing at endline if netiher *i* nor *j* is selected to participate. Similarly, $\delta_1 + \delta_2$ identifies the effect if only one participates, while $\delta_1 + \delta_2 + \delta_3$ indicates the effect of both are participants.

Panel A is restricted to those who were not OR friends at baseline and thus shows the effect on the formation of new links. In Columns (1) and (2), we see that all pairs are more likely to be friends in T2 schools than in C. Further, pairs in which exactly one is selected and pairs in which both are both significantly more likely to be friends at endline. In Columns (3) and (4), however, we see that only pairs in which neither or only one participate are significantly more likely to be OR friends. Taken together, the results in Panel A suggest that, among those who were not friends at baseline, the program led to more OR

friendships among pairs of non-participants and more AND friendships among pairs of participants. This suggests that partitioning did not result from selective formation of new friendships.

Panels B and C present results for those who were friends at baseline, and suggests that the network partitionig results are driven by these groups. In all specifications, coefficients δ_1 and δ_3 are positive while δ_2 is negative. Panel B suggests that the primary effect of the program for OR friends was to reinforce these friendships for pairs of non-participants while also changing many of these to AND friendships for both pairs of non-participants and pairs of participants. In Panel C, we see that the program significantly affected the probability of being both AND and OR friends at endline among pairs who were AND friends at baseline.

4. Program Effects on Aspirations and Attitudes

Intent to Treat Estimates

After documenting the effect of the program on network measures, we measure the effect of the program on attitudes and aspirations. These outcomes are constructed as mean zero, variance one indices as described above. We then estimate the following:

10) $y_{is1} = \beta_0 + \beta_1 T1_s + \beta_2 T2_s + \beta_3 y_{is0} + \epsilon_{is1}$

The parameter β_1 identifies the average effect of the program on all students in T1 schools, while β_2 identifies the average effect of the program on all students in T2. Baseline outcomes y_{is0} are included for precision. We cluster all standard errors by school.

Table 11 presents Intention to Treat estimates for aspirations, expectations, and attitudes as measured by our four endline indices. These results are pooled for all students, including boys, participant girls, and non-participant girls. Surpringly, and counter to our priors, we see negative point estimates on all outcomes in T1, and on three of four outcomes in T2. The strongest effects appear to be on the self confidence measure, as all specifications show significantly negative effects of between 0.3 and 0.4 standard deviations in both T1 and T2 schools. We also see significant negative effects on educational aspirations in T1 schools in Column (2) when controlling for baseline outcomes.

Heterogeneous Treatment Effects

Finally, we estimate heterogeneous treatment effects to allow for disaggregation of effects on three groups that may be affected differently by the program: participant girls, non-participant girls, and boys. The baseline individual-level analysis showed that girls and boys in these schools are different along a

number of dimensions. Additionally, as the program was targeted specifically at girls, there is reason to believe that its effect may be different on girls and boys.

Further, both the individual-level and link-level analyses presented above provide strong evidence that elected and non-elected girls are different among multiple dimensions. Additionally, only 13 girls in each school actually participate in the program, and thus the effects on participants and non-participants may be quite different. To look at heterogeneity, we present estimates of specifications in Equation 11.

11) $y_{is1} = \beta_0 + \beta_1 T1_s + \beta_2 T1_s * Girl_{is} + \beta_3 T1_s * Girl_{is} * E_{is} + \beta_4 T2_s + \beta_5 T2_s * Girl_{is} + \beta_6 T2_s * Girl_{is} * P_{is} + \beta_7 E_{is} + \beta_8 y_{is0} + \epsilon_{is1}$

In Equation 11, the parameters β_1 and β_4 identify the effect of the program on boys in T1 and T2 schools, respectively. $\beta_1 + \beta_2$ and $\beta_4 + \beta_5$ identify the effect on non-participant girls in these schools, while $\beta_1 + \beta_2 + \beta_3$ and $\beta_4 + \beta_5 + \beta_6$ identify the effect on participant girls.

Results are presented in Table 12. Interestingly, we see relatively little evidence of impacts on participant girls. Career aspirations and expectations appear to be negatively impacted in T1 schools, but this result is only marginally significant. However, it is subsantively large at approximately 0.3 standard devisions. Additionally, self-confidence among randomly-selected participants in T2 appears to be negatively impacted in T2 schools.

Non-participant girls have significantly lower educational aspirations at endline in T1 schools, as shown by the p-value of the test on the coefficient for $T1_s + T1_s * Girl_{is}$.We further see evidence of negative effects on non-participant girls' self-confidence in T1, but note that there is substantially less evidence for negative effects on self-confidence for non-participant girls in T2. We interpret this as an effect of the selection mechanism in T1: girls who were not chosen by election lose self confidence, and this effect is re-emphasized every time the Bal Sabha meets without their participation.

Finally, and most starkly, the estimates in Table 12 suggest that the largest impacts of the program may be on boys' self confidence. Average boys' self-confidence is approximately 0.40 standard deviations lower in T1 schools and 0.476 standard deviations lower in T2 schools. Boys, who were not the target of the program, ended up with statistically significant and quantitatively meaningfully lower self confidence in schools that received the girls-targeted program.

5. Conclusion

This paper examines network formation and outcomes due to an after-school program in rural India, using extensive panel network data, combined with novel randomized assignment to selection regime and program participation. Several important empirical findings emerge from this analysis.

We find substantial evidence of friendship network sorting as well as of network segregation between those who are selected and those who are not due to program exclusion. Network segregation occurs through two different channels depending on the selection regime. Selective exclusion partitions networks during the selection process itself – based on elections or on characteristic-based eligibility (e.g. gender). Two elected girls are 24.9 percentage points more likely to be friends at baseline than two non-elected girls and 15.0 percentage points more likely to be friends than if only one is elected. When exclusion is random, network segregation is due to program participation (or non-participation). Pairs of participants being more likely to be friends at endline than pairs of non-participants, and pairs in which one is a participant and the other not being less likely to be friends than either group.

In addition, we find negative spillovers when girls are selected by popular vote, translating into lower levels of self-confidence among girls who were not elected. Non-elected girls in schools running the program have a self-confidence index 0.37 standard deviations lower than those in the control group, suggesting a discouragement effect from not being elected. We do not find these negative spillovers when exclusion/participation is determined randomly. Negative spillovers affect boys' self-confidence in both selection regimes.

Our findings have important implications for the estimation of peer effects and the design of appropriate rules for assigning individuals to social programs in a wide range of areas. The evidence in this paper calls for caution in expanding this type of education program based on the endogenous exclusion of a significant portion of the school population. "Ensuring *inclusive* and equitable quality education", "promoting sustained, *inclusive* and sustainable economic growth", and "promoting peaceful and *inclusive* societies" are 3 of the 17 Sustainable Development Goals (UN 2015). Social inclusion encompasses a sense of belonging, of integration to the reference group (Shortall, 2008). Self-confidence and social capital are significant contributors to social inclusion (Bailey 2005, Fiorina 1999). And, while social programs are designed to empower and bring opportunities, this paper shows that selective social programs may in some cases, undermine the self-confidence and social capital of the excluded segments of the population.

References

Bandiera, Oriana, and Imran Rasul. 2006. "Social Networks and Technology Adoption in Northern Mozambique." *Economic Journal* 116(514): 869-902.

Bandura, Albert, Claudio Barbaranelli, Gian Vittorio Caprara, and Concetta Pastorelli. 2001. "Self-Efficacy Beliefs as Shapers of Children's Aspirations and Career Trajectories" *Child Development* 72(1): 187-206.

Banerjee, Abhijit, Arun G. Chandrasekhar, Esther Duflo, and Matthew O. Jackson. 2012. "The Diffusion of Microfinance." Unpublished Working Paper.

Beaman, Lori, Raghabendra Chattopadhyay, Esther Duflo, and Petia Topalova. 2009. "Powerful Women: Femael Leadership and Gender Bias." *Quarterly Journal of Economics* 124(4): 1497-1540.

Beaman, Lori, and Jeremy Magruder. 2012. "Who Gets the Job Referral? Evidence from a Socail Networks Experiment." *American Economic Review* 102(7): 3574-3593.

Bernard, Tanguy, Stefan Dercon, and Alemayehu Seyoum Taffesse. 2011. "Beyond Fatalism—An Empirical Exploration of Self-Efficacy and Aspirations Failure in Ethiopia." Unpublished Working Paper.

Burgess, Simon, Sanderson, Eleanor and Umana-Aponte, Marcela, (2011), School ties: An analysis of homophily in an adolescent friendship network, The Centre for Market and Public Organisation, Department of Economics, University of Bristol, UK.

Comola, Margherita, and Silvia Prina. 2014. "Do Interventions Change the Nework? A Dynamic Peer Effect Model Accounting for Network Changes." Unpublished Working Paper.

Conley, Timonthy G. and Christopher R. Udry. 2012. "Learning about a New Technology: Pineapple in Ghana." *American Economic Review*: 35-69.

Currarini, Sergio, Matthew O. Jackson, and Paolo Pin. 2009. "An Economic Model of Friendship: Homophily, Minorities, and Segregation." *Econometrica* 77(4): 1003-1045.

De Giorgi , Giacomo, Michele Pellizzarri, and Silvia Radaelli. 2010. "Identification of Social Interactions through Partially Overlapping Peer Groups." *American Economic Journal: Applied Economics* 2: 241-275.

Dercon, Stefan, and Abhijeet Singh. 2013. "From Nutrition to Aspirations and Self-Efficacy: Gender Bias Over Time among Children in Four Countries." World Development 45: 31-50.

French, D., Jansen, E., Riansari, M. and Setiono, K. (2003). Friendships of Indonesian children: Adjustment of children who differ in friendship presence and similarity between mutual friends, Social Development 12(4): 605–621.

Hamm, Jill V. “Do birds of a feather flock together? The variable bases for African American, Asian American, and European American adolescents' selection of similar friends.” Developmental Psychology, Vol 36(2), Mar 2000, 209-219.

Heckman, James and Yona Rubinstein. 2001. "The Importance of Noncognitive Skills: Lessons from the GED Testing Program," American Economic Review, vol. 91(2): 145-149.

Kandel, Denise B.. 1978. “Homophily, Selection, and Socialization in Adolescent Friendships”. American Journal of Sociology 84 (2). University of Chicago Press: 427–36.

Kosec, Katrina, Madeeha Hameed, and Stephanie Hausladen. 2012. “Aspirations in Rural Pakistan.” Unpublished Working Paper.

Manski, Charles F. 1993. “Identification of Endogenous Social Effects: The Reflection Problem.” *Review of Economic Studies* 60(3): 531-542.

Miguel, Edward, and Michael Kremer. 2004. “Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities.” *Econometrics* 72(1): 159-217.

Ngatia, Muthoni. 2011. “Social Interactions and Individual Reproductive Decisions.” Unpublished Working Paper.

Nguyen, Trang. 2008. “Information, Role Models, and Perceived Returns to Education: Experimental Evidence from Madagascar.” Unpublished Working Paper.

Oster, Emily, and Rebecca Thornton. 2012. “Determinants of Technology Adoption: Peer Effects in Menstrual Cup Take-up.” *Journal of the European Economic Association* 10(6): 1263-1293.

Sacerdote, Bruce. 2001. “Peer Effects with Random Assignment: Results for Dartmouth Roommates.” *Quarterly Journal of Economics* 116(2): 681-704.

United Nations. 2015. “Transforming our world: the 2030 Agenda for Sustainable Development” Resolution adopted by the General Assembly on 25 September 2015, A/RES/70/1.

Vasilaky, Kathryn, and Kenneth L. Leonard. 2014. “As Good as the Networks They Keep?: Improving Farmers’ Social Networks via Randomized Information Exchange in Rural Uganda.” Unpublished Working Paper.

Table 1: Baseline Sample

	Girls (N=1414)	Boys (N=1196)	Difference	P-value of F-test of Equalty
	(1)	(2)	(3)	(4)
Standard	6.931	7.019	0.088	0.086
	(0.001)	(0.001)		
Age	12.325	12.558	0.232	0.021
	(0.004)	(0.008)		
Scheduled Caste	0.255	0.327	0.072	0.030
	(0.001)	(0.002)		
Scheduled Tribe	0.123	0.262	0.139	0.000
	(0.001)	(0.002)		
Other Backwards Caste	0.445	0.284	-0.161	0.000
	(0.002)	(0.001)		
Enrolled Previous Year	0.840	0.846	0.006	0.833
	(0.001)	(0.001)		
Owns TV	0.868	0.779	-0.089	0.013
	(0.001)	(0.001)		
Father Attended School	0.831	0.694	-0.137	0.000
	(0.001)	(0.002)		
Mother Attended School	0.562	0.394	-0.168	0.000
	(0.001)	(0.002)		
Education Index	-0.180	0.221	0.401	0.001
	(0.010)	(0.007)		
Career Index	-0.129	0.151	0.281	0.012
	(0.006)	(0.008)		
Self-confidence Index	-0.009	0.008	0.017	0.856
	(0.008)	(0.002)		
Gender roles Index	0.112	-0.135	-0.248	0.020
	(0.011)	(0.008)		
Number of Friends (OR)	15.761	--	--	--
	(2.353)	--		
Number of Friends (AND)	7.815	--	--	--
	(0.640)	--		

Notes: Robust standard errors in parentheses, clustered by school.
*** p<0.01, ** p<0.05, * p<0.1
Individuals *i* and *j* are "OR" friends if (at least) one names the other as a friend.
Individuals *i* and *j* are "AND" friends if they both name each other as friends.

Table 2: Baseline Balance - School-level Randomization

	Girls				Boys			
	Control	Regression coefficent on T1	Regression coefficent on T2	P-value of Joint test	Control	Regression coefficent on T1	Regression coefficent on T2	P-value of Joint test
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Standard	6.934	0.069	-0.073	0.193	6.997	-0.030	0.114	0.114
	(0.016)	(0.047)	(0.070)		(0.053)	(0.059)	(0.081)	
Age	12.458	-0.217	-0.212	0.281	12.713	-0.339	-0.081	0.170
	(0.112)	(0.152)	(0.146)		(0.160)	(0.197)	(0.205)	
Scheduled Caste	0.293	-0.097	-0.030	0.550	0.399	-0.136	-0.068	0.406
	(0.076)	(0.095)	(0.100)		(0.066)	(0.100)	(0.115)	
Scheduled Tribe	0.070	0.068	0.103	0.218	0.201	0.062	0.125	0.524
	(0.026)	(0.045)	(0.086)		(0.050)	(0.070)	(0.133)	
Other Backwards Caste	0.430	0.016	0.032	0.957	0.282	0.007	-0.002	0.991
	(0.083)	(0.099)	(0.108)		(0.050)	(0.070)	(0.081)	
Enrolled Previous Year	0.829	0.034	0.002	0.860	0.843	-0.002	0.016	0.931
	(0.050)	(0.069)	(0.078)		(0.041)	(0.058)	(0.055)	
Owns TV	0.904	-0.080	-0.035	0.335	0.788	-0.002	-0.031	0.964
	(0.022)	(0.056)	(0.046)		(0.048)	(0.062)	(0.115)	
Father Attended School	0.848	0.005	-0.059	0.461	0.648	0.077	0.068	0.748
	(0.040)	(0.058)	(0.056)		(0.096)	(0.101)	(0.107)	
Mother Attended School	0.561	0.065	-0.065	0.310	0.337	0.136	0.028	0.183
	(0.055)	(0.076)	(0.085)		(0.070)	(0.080)	(0.107)	
Education Index	-0.212	-0.017	0.118	0.785	0.315	-0.106	-0.195	0.681
	(0.173)	(0.262)	(0.213)		(0.134)	(0.185)	(0.228)	
Career Index	-0.033	0.060	-0.362	0.034	0.353	-0.305	-0.311	0.311
	(0.119)	(0.132)	(0.184)		(0.186)	(0.210)	(0.204)	
Self-confidence Index	0.075	-0.317	0.053	0.107	0.051	-0.093	-0.023	0.655
	(0.148)	(0.206)	(0.173)		(0.076)	(0.105)	(0.106)	
Gender Roles Index	0.194	-0.077	-0.178	0.770	0.061	-0.203	-0.440	0.235
	(0.116)	(0.221)	(0.254)		(0.167)	(0.188)	(0.251)	
Number of Friends (OR)	16.581	-1.989	-0.721	0.863				
	(2.478)	(3.676)	(3.628)					
Number of Friends (AND)	6.793	0.933	2.385	0.479				
	(0.769)	(1.440)	(2.072)					
Proportion of Friends Elected (OR)	0.275	0.097	0.045	0.648				
	(0.071)	(0.104)	(0.081)					
Proportion of Friends Elected (AND)	0.284	0.095	0.030	0.639				
	(0.066)	(0.101)	(0.077)					

Notes: Robust standard errors in parentheses, clustered by school.

*** p<0.01, ** p<0.05, * p<0.1

Individuals *i* and *j* are "OR" friends if (at least) one names the other as a friend.

Individuals *i* and *j* are "AND" friends if they both name each other as friends.

Table 3: Baseline Balance - Individual Randomization to Program among T2 Girls

	Not Selected	Regression coefficient on Selected	P-value of Joint test
	(1)	(2)	(3)
Standard	6.976	-0.032	0.861
	(0.037)	(0.175)	
Age	12.460	-0.277	0.163
	(0.083)	(0.182)	
Scheduled Caste	0.295	-0.009	0.803
	(0.074)	(0.036)	
Scheduled Tribe	0.249	-0.050	0.110
	(0.103)	(0.028)	
Other Backwards Caste	0.371	0.065	0.232
	(0.069)	(0.051)	
Enrolled Previous Year	0.850	-0.043	0.532
	(0.041)	(0.066)	
Owns TV	0.810	0.067	0.265
	(0.076)	(0.056)	
Father Attended School	0.775	-0.100	0.046
	(0.039)	(0.043)	
Mother Attended School	0.420	0.110	0.220
	(0.067)	(0.084)	
Education Index	0.024	-0.148	0.190
	(0.127)	(0.105)	
Career Index	-0.161	-0.290	0.126
	(0.089)	(0.172)	
Self-confidence Index	0.084	0.014	0.920
	(0.070)	(0.139)	
Gender Roles Index	-0.171	0.127	0.265
	(0.197)	(0.107)	
Number of Friends (OR)	15.938	-0.266	0.859
	(2.896)	(1.448)	
Number of Friends (AND)	8.911	0.913	0.446
	(2.105)	(1.145)	
Proportion of Friends Elected (OR)	0.298	0.072	0.101
	(0.037)	(0.040)	
Proportion of Friends Elected (AND)	0.281	0.107	0.014
	(0.037)	(0.035)	

Notes: Robust standard errors in parentheses, clustered by school.

*** p<0.01, ** p<0.05, * p<0.1

Individuals *i* and *j* are "OR" friends if (at least) one names the other as a friend.

Individuals *i* and *j* are "AND" friends if they both name each other as friends.

Table 4: Survey Attrition

Panel A: School randomization	Girls				Boys			
Present for:	Control	Regression coefficent on T1	Regression coefficent on T2	P-value of Joint Test on T1 and T2	Control	Regression coefficent on T1	Regression coefficent on T2	P-value of Joint Test on T1 and T2
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Baseline survey	0.688	0.099	0.072	0.240	0.784	-0.011	-0.140	0.041
	(0.042)	(0.059)	(0.057)		(0.050)	(0.062)	(0.064)	
Baseline network survey	0.710	-0.015	0.033	0.573	0.661	0.075	-0.012	0.182
	(0.033)	(0.044)	(0.048)		(0.059)	(0.066)	(0.069)	
Endline survey	0.714	0.071	0.031	0.276	0.753	0.049	-0.054	0.172
	(0.030)	(0.043)	(0.048)		(0.033)	(0.045)	(0.056)	
Endline network survey	0.712	0.063	0.033	0.391	0.779	0.016	-0.077	0.397
	(0.031)	(0.046)	(0.042)		(0.023)	(0.039)	(0.064)	

Panel B: Randomization to participate among T2 Girls

Present for:	Not Selected	Regrssion coefficient on Selected	P-value of Test of Selected
	(1)	(2)	(3)
Baseline survey	0.735	0.092	0.045
	(0.042)	(0.040)	
Baseline network survey	0.704	0.139	0.023
	(0.043)	(0.051)	
Endline survey	0.701	0.158	0.002
	(0.042)	(0.038)	
Endline network survey	0.701	0.158	0.015
	(0.040)	(0.053)	

Notes: Robust standard errors in parentheses, clustered by school.
*** p<0.01, ** p<0.05, * p<0.1
Sample for Panel A is all enrolled students. N = 1426 for girls and N = 1229 for boys.
Sample for Panel B is all enrolled girls in T2 schools. N = 451.

Table 5: Baseline Endogenously-formed *OR* and *AND* Networks

	Absolute Value of Distance between Students Values		
Network Definition	OR	AND	Observations
	(1)	(2)	(3)
Standard	-0.102***	-0.154***	27,418
	(0.012)	(0.018)	
Age	-0.019***	-0.031***	27,298
	(0.007)	(0.010)	
Scheduled Caste	0.001	0.000	27,210
	(0.020)	(0.030)	
Scheduled Tribe	-0.067	-0.057	27,210
	(0.061)	(0.042)	
Other Backwards Caste	-0.010	-0.014	27,210
	(0.015)	(0.019)	
Enrolled Previous Year	-0.092***	-0.117***	24,636
	(0.033)	(0.040)	
Family owns TV	-0.040	-0.048	14,428
	(0.032)	(0.060)	
Father Attended School	-0.005	-0.010	13,774
	(0.024)	(0.038)	
Mother Attended School	-0.019	-0.033	13,820
	(0.024)	(0.027)	
Education Index	-0.005	-0.027	13,178
	(0.012)	(0.017)	
Career Index	0.004	-0.013	18,070
	(0.014)	(0.018)	
Self-confidence Index	-0.009	-0.023*	18,868
	(0.008)	(0.012)	
Gender Roles Index	-0.031**	-0.049***	18,424
	(0.016)	(0.017)	
Number of Friends (OR)	-0.004***	-0.008***	27,418
	(0.001)	(0.002)	
Number of Friends (AND)	0.002	-0.008***	27,418
	(0.001)	(0.003)	

Notes: Robust standard errors in parentheses, clustered by school.
*** p<0.01, ** p<0.05, * p<0.1
Dependent variable is existence of friendship at Baseline under the appropriate network definition.
Reported values are regression coefficient of distance between students' values for each variable.

Table 6: Baseline Characteristics of Girls Elected and Not Elected

	Elected (N=374)	Not Elected (N=1040)	Difference	P-value of F-test of Equalty
	(1)	(2)	(3)	(4)
Standard	7.112	6.866	0.246	0.000
	(0.058)	(0.030)	(0.059)	
Age	12.497	12.237	0.260	0.030
	(0.101)	(0.061)	(0.103)	
Scheduled Caste	0.273	0.246	0.026	0.569
	(0.058)	(0.036)	(0.044)	
Scheduled Tribe	0.094	0.133	-0.039	0.127
	(0.038)	(0.033)	(0.026)	
Other Backwards Caste	0.422	0.455	-0.033	0.546
	(0.056)	(0.043)	(0.050)	
Enrolled Previous Year	0.853	0.832	0.021	0.664
	(0.039)	(0.035)	(0.040)	
Owns TV	0.869	0.867	0.002	0.945
	(0.029)	(0.027)	(0.030)	
Father Attended School	0.828	0.833	-0.005	0.845
	(0.032)	(0.024)	(0.026)	
Mother Attended School	0.538	0.575	-0.037	0.400
	(0.050)	(0.032)	(0.043)	
Education Index	-0.033	-0.261	0.227	0.019
	(0.122)	(0.102)	(0.091)	
Career Index	-0.072	-0.160	0.088	0.259
	(0.083)	(0.087)	(0.077)	
Self-confidence Index	0.006	-0.017	0.023	0.814
	(0.082)	(0.109)	(0.097)	
Gender Roles Index	0.109	0.114	-0.005	0.952
	(0.120)	(0.106)	(0.083)	
Number of Friends (OR)	15.251	15.951	-0.700	0.586
	(1.404)	(1.693)	(1.271)	
Number of Friends (AND)	8.698	7.486	1.212	0.087
	(1.011)	(0.797)	(0.685)	
Proportion of Friends Elected (OR)	0.424	0.273	0.151	0.001
	(0.041)	(0.037)	(0.039)	
Proportion of Friends Elected (AND)	0.460	0.262	0.197	0.000
	(0.039)	(0.034)	(0.038)	

Notes: Robust standard errors in parentheses, clustered by school.

*** p<0.01, ** p<0.05, * p<0.1

Individuals *i* and *j* are "OR" friends if (at least) one names the other as a friend.

Individuals *i* and *j* are "AND" friends if they both name each other as friends.

Table 7: Baseline Network Links and Election Results

Network Definition	OR	AND
	(1)	(2)
Elected (OR)	0.080	0.099*
	(0.049)	(0.052)
Elected (AND)	0.073***	0.150***
	(0.022)	(0.041)
Constant	0.746***	0.336***
	(0.062)	(0.063)
Observations	27,418	27,418
R-squared	0.015	0.023
Mean of Dep var in Control	0.028	0.003
P-value of Elected (OR) + Elected (AND)	0.028	0.003

Notes: Robust standard errors in parentheses, clustered by school.

*** p<0.01, ** p<0.05, * p<0.1

Dependent variable is existence of friendship at Baseline, under appropriate definition.

Individuals *i* and *j* are "OR" friends if (at least) one names the other as a friend.

Individuals *i* and *j* are "AND" friends if they both name each other as friends.

Table 8: ITT Program Effects on Endline Network Formation

Network Definition	OR				AND			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
T1	0.121	0.068	0.063	0.013	0.061	-0.005	-0.015	-0.067
	(0.092)	(0.067)	(0.060)	(0.058)	(0.078)	(0.052)	(0.044)	(0.060)
T2	0.189**	0.122**	0.131**	0.078	0.180**	0.096	0.094	0.053
	(0.076)	(0.056)	(0.056)	(0.050)	(0.076)	(0.058)	(0.059)	(0.065)
Friends at Baseline (OR)		0.192***	0.158***	0.154***		0.188***	0.160***	0.157***
		(0.015)	(0.015)	(0.014)		(0.019)	(0.020)	(0.020)
Friends at Baseline (AND)		0.162***	0.121***	0.120***		0.244***	0.184***	0.184***
		(0.033)	(0.031)	(0.029)		(0.030)	(0.033)	(0.032)
Constant	0.631***	0.446***	0.370***	1.105*	0.325***	0.118***	0.084**	0.923
	(0.071)	(0.063)	(0.067)	(0.586)	(0.065)	(0.040)	(0.039)	(0.703)
Baseline Network Controls Included	NO	NO	YES	YES	NO	NO	YES	YES
School Size Controls Included	NO	NO	NO	YES	NO	NO	NO	YES
Observations	15,578	15,578	15,578	15,578	15,578	15,578	15,578	15,578
R-squared	0.033	0.114	0.137	0.144	0.024	0.133	0.161	0.165
Mean of Dep var in Control	0.631	0.631	0.631	0.631	0.325	0.325	0.325	0.325
P-value of Test of T1 = T2	0.296	0.336	0.209	0.239	0.056	0.060	0.052	0.049

Notes: Robust standard errors in parentheses, clustered by school.

*** p<0.01, ** p<0.05, * p<0.1

Dependent variable is existence of friendship at Endline.

Baseline Network Controls include answers to all baseline network survey questions.

School Size Controls include linear, quadratic, and cubic in number of students enrolled in the school at beginning of school year.

Individuals *i* and *j* are "OR" friends if (at least) one names the other as a friend.

Individuals *i* and *j* are "AND" friends if they both name each other as friends.

Table 9: Disaggregated ITT Program Effects on Endline Network Formation

Network Definition	OR				AND			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Elected (OR)	0.100**	0.079**	0.066**	0.053*	0.121***	0.095***	0.082***	0.074***
	(0.038)	(0.029)	(0.027)	(0.027)	(0.036)	(0.027)	(0.024)	(0.024)
Elected (AND)	0.128***	0.093***	0.075**	0.064*	0.179***	0.134***	0.114***	0.107***
	(0.032)	(0.027)	(0.032)	(0.034)	(0.034)	(0.028)	(0.030)	(0.030)
T1	0.099	0.045	0.045	0.007	0.036	-0.031	-0.033	-0.069
	(0.096)	(0.071)	(0.066)	(0.069)	(0.065)	(0.039)	(0.035)	(0.057)
T1 * Elected (OR)	-0.003	0.015	0.014	0.021	-0.010	0.011	0.008	0.015
	(0.055)	(0.051)	(0.049)	(0.049)	(0.042)	(0.035)	(0.034)	(0.031)
T1 * Elected (AND)	-0.036	-0.023	-0.022	-0.018	-0.034	-0.023	-0.023	-0.019
	(0.062)	(0.050)	(0.051)	(0.048)	(0.080)	(0.057)	(0.061)	(0.057)
T2	0.190**	0.131**	0.150**	0.109*	0.167*	0.094	0.103	0.082
	(0.071)	(0.055)	(0.058)	(0.055)	(0.082)	(0.068)	(0.070)	(0.077)
T2 * Selected (OR)	-0.065**	-0.059**	-0.067***	-0.063***	-0.068	-0.062	-0.068	-0.068
	(0.025)	(0.022)	(0.021)	(0.021)	(0.046)	(0.043)	(0.042)	(0.042)
T2 * Selected (AND)	0.043*	0.025	0.013	0.019	0.109**	0.086*	0.069**	0.072**
	(0.024)	(0.021)	(0.021)	(0.020)	(0.050)	(0.042)	(0.033)	(0.034)
Friends at Baseline (OR)		0.185***	0.155***	0.152***		0.180***	0.157***	0.155***
		(0.013)	(0.014)	(0.013)		(0.018)	(0.019)	(0.019)
Friends at Baseline (AND)		0.152***	0.120***	0.119***		0.230***	0.183***	0.183***
		(0.031)	(0.029)	(0.028)		(0.030)	(0.031)	(0.031)
Constant	0.590***	0.422***	0.357***	0.875	0.275***	0.089**	0.070*	0.570
	(0.068)	(0.059)	(0.064)	(0.597)	(0.056)	(0.033)	(0.036)	(0.712)
Baseline Network Controls Include	NO	NO	YES	YES	NO	NO	YES	YES
School Size Controls Included	NO	NO	NO	YES	NO	NO	NO	YES
Observations	15,578	15,578	15,578	15,578	15,578	15,578	15,578	15,578
R-squared	0.057	0.129	0.147	0.151	0.058	0.155	0.176	0.177
Mean of Dep var in Control	0.631	0.631	0.631	0.631	0.325	0.325	0.325	0.325
P-value of Test of T1 + T1 * Elected (OR)	0.165	0.286	0.268	0.635	0.665	0.653	0.569	0.389
Elected (OR) + T1 * Elected (AND)	0.368	0.499	0.495	0.878	0.923	0.565	0.537	0.421
P-value of Test of T2 + T2 * Selected (OR)	0.045	0.145	0.101	0.335	0.101	0.535	0.508	0.828
Selected (OR) + T2 * Selected (AND)	0.007	0.047	0.061	0.207	0.001	0.029	0.021	0.097

Notes: Robust standard errors in parentheses, clustered by school.

*** p<0.01, ** p<0.05, * p<0.1

Dependent variable is existence of friendship at Endline.

Baseline Network Controls include answer to all baseline network survey questions, as well as interactions between (IN) and (OUT).

School Size Controls include linear, quadratic, and cubic in number of students enrolled in the school at beginning of school year.

Individuals *i* and *j* are "OR" friends if (at least) one names the other as a friend.

Individuals *i* and *j* are "AND" friends if they both name each other as friends.

Table 10: Disaggregated Program ITT Effects on Network Formation

Panel A: Not Friends at Baseline

Network Definition	OR			AND		
	(1)	(2)	(3)	(4)	(5)	(6)
T2	0.255***	0.233***	0.127*	0.076	0.069	-0.002
	(0.076)	(0.067)	(0.064)	(0.051)	(0.045)	(0.043)
T2 * P (OR)	-0.056	-0.094*	-0.095*	0.056	0.035	0.033
	(0.047)	(0.052)	(0.048)	(0.041)	(0.034)	(0.032)
T2 * P (AND)	-0.074	-0.082	-0.068	0.126	0.113	0.114
	(0.102)	(0.110)	(0.110)	(0.114)	(0.089)	(0.086)
Constant	0.420***	0.319***	1.609	0.106***	0.068**	1.348*
	(0.058)	(0.050)	(1.003)	(0.032)	(0.024)	(0.657)
Baseline Network Controls Included	NO	YES	YES	NO	YES	YES
School Size Controls Included	NO	NO	YES	NO	NO	YES
Observations	2,544	2,544	2,544	2,544	2,544	2,544
R-squared	0.034	0.091	0.106	0.025	0.101	0.115
Mean of Dep var in Control	0.631	0.631	0.631	0.325	0.325	0.325
P-value of Test of T2 + T2 * Participant (OR)	0.008	0.086	0.680	0.026	0.080	0.588
P-value of Test of T2 + T2 * Participant (OR) + T2 * Participant (AND)	0.345	0.669	0.781	0.022	0.014	0.083

Panel B: "OR" Friends but not "AND" Friends at Baseline

Network Definition	OR			AND		
	(1)	(2)	(3)	(4)	(5)	(6)
T2	0.184**	0.200**	0.121*	0.137	0.153*	0.093
	(0.080)	(0.078)	(0.066)	(0.085)	(0.085)	(0.090)
T2 * P (OR)	-0.062	-0.081**	-0.076**	-0.071	-0.081	-0.077
	(0.036)	(0.034)	(0.035)	(0.066)	(0.058)	(0.061)
T2 * P (AND)	0.048	0.005	0.010	0.179**	0.138***	0.139***
	(0.067)	(0.052)	(0.056)	(0.072)	(0.048)	(0.045)
Constant	0.623***	0.510***	1.880*	0.288***	0.191***	1.736
	(0.065)	(0.076)	(1.085)	(0.054)	(0.049)	(1.221)
Baseline Network Controls Included	NO	YES	YES	NO	YES	YES
School Size Controls Included	NO	NO	YES	NO	NO	YES
Observations	4,598	4,598	4,598	4,598	4,598	4,598
R-squared	0.028	0.083	0.096	0.019	0.063	0.074
Mean of Dep var in Control	0.631	0.631	0.631	0.325	0.325	0.325
P-value of Test of T2 + T2 * Participant (OR)	0.119	0.080	0.409	0.428	0.318	0.816
P-value of Test of T2 + T2 * Participant (OR) + T2 * Participant (AND)	0.120	0.184	0.542	0.029	0.009	0.044

Panel C: "AND" Friends at Baseline

Network Definition	OR			AND		
	(1)	(2)	(3)	(4)	(5)	(6)
T2	0.047	0.044	0.050	0.063	0.046	0.067
	(0.055)	(0.048)	(0.051)	(0.085)	(0.083)	(0.089)
T2 * P (OR)	-0.023	-0.021	-0.025	-0.040	-0.054	-0.061
	(0.028)	(0.028)	(0.027)	(0.042)	(0.044)	(0.041)
T2 * P (AND)	0.073**	0.078***	0.082**	0.082	0.075*	0.081*
	(0.034)	(0.027)	(0.029)	(0.050)	(0.039)	(0.041)
Constant	0.845***	0.803***	0.948*	0.587***	0.522***	0.721
	(0.030)	(0.050)	(0.538)	(0.051)	(0.081)	(0.951)
Baseline Network Controls Included	NO	YES	YES	NO	YES	YES
School Size Controls Included	NO	NO	YES	NO	NO	YES
Observations	4,442	4,442	4,442	4,442	4,442	4,442
R-squared	0.007	0.045	0.049	0.005	0.079	0.085
Mean of Dep var in Control	0.631	0.631	0.631	0.325	0.325	0.325
P-value of Test of T2 + T2 * Participant (OR)	0.602	0.591	0.609	0.736	0.909	0.936
P-value of Test of T2 + T2 * Participant (OR) + T2 * Participant (AND)	0.013	0.011	0.017	0.083	0.184	0.167

Notes: Robust standard errors in parentheses, clustered by school.

*** p<0.01, ** p<0.05, * p<0.1

Sample restricted to girls in Control and T2 schools.

Dependent variable is existence of friendship at Endline.

Baseline Network Controls include answer to all baseline network survey questions, as well as interactions between (IN) and (OUT).

School Size Controls include linear, quadratic, and cubic in number of students enrolled in the school at beginning of school year.

Individuals *i* and *j* are "OR" friends if (at least) one names the other as a friend.

Individuals *i* and *j* are "AND" friends if they both name each other as friends.

Table 11: Program ITT Effects on Endline Attitudes

	Education Index		Career Index		Self-Confidence Index		Gender Roles Index	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
T1	-0.221	-0.182*	-0.113	-0.111	-0.351**	-0.343**	-0.071	-0.062
	(0.142)	(0.096)	(0.089)	(0.089)	(0.147)	(0.149)	(0.146)	(0.146)
T2	-0.091	-0.056	0.007	0.015	-0.325**	-0.327**	-0.113	-0.091
	(0.153)	(0.096)	(0.115)	(0.117)	(0.157)	(0.156)	(0.169)	(0.164)
Baseline Response		0.530***		0.020		0.032*		0.083*
		(0.037)		(0.033)		(0.018)		(0.042)
Constant	0.141	0.102	0.041	0.038	0.258**	0.255**	0.080	0.069
	(0.091)	(0.072)	(0.046)	(0.047)	(0.106)	(0.106)	(0.077)	(0.080)
Observations	1,297	1,297	1,552	1,552	1,585	1,585	1,568	1,568
R-squared	0.009	0.290	0.003	0.003	0.026	0.027	0.002	0.009
Mean of dep var in control	0.141	0.141	0.041	0.041	0.258	0.258	0.080	0.080
P-value of test of T1 = T2	0.438	0.173	0.363	0.345	0.871	0.915	0.834	0.880

Notes: Robust standard errors in parentheses, clustered by school.
*** p<0.01, ** p<0.05, * p<0.1
Dependent variable is the first principal component of responses to relevant questions on the endline questionnaire.

Table 12 -- Intention to Treat Estimates

Dependent var: Endline Response to	Educational Aspirations / Expectations Index		Career Aspirations / Expectations Index		Self-Confidence Index		Gender Roles Index	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
T1	-0.148	-0.065	-0.057	-0.055	-0.402**	-0.398**	-0.031	-0.016
	(0.149)	(0.099)	(0.142)	(0.143)	(0.159)	(0.159)	(0.168)	(0.171)
T1 * Girl	-0.247	-0.270	-0.067	-0.070	0.030	0.038	0.007	0.003
	(0.267)	(0.164)	(0.174)	(0.174)	(0.180)	(0.178)	(0.132)	(0.139)
T1 * Elected	0.195	0.100	-0.173	-0.174	0.147	0.142	-0.216	-0.235
	(0.182)	(0.135)	(0.206)	(0.205)	(0.249)	(0.247)	(0.178)	(0.178)
T2	-0.156	0.004	0.110	0.113	-0.476**	-0.476**	-0.136	-0.099
	(0.210)	(0.119)	(0.184)	(0.183)	(0.218)	(0.217)	(0.221)	(0.220)
T2 * Girl	0.191	-0.058	-0.197	-0.196	0.313	0.311	0.103	0.075
	(0.285)	(0.161)	(0.170)	(0.171)	(0.208)	(0.208)	(0.231)	(0.227)
T2 * Participant	-0.190	-0.142	0.101	0.102	-0.154*	-0.153*	-0.199*	-0.191**
	(0.157)	(0.106)	(0.137)	(0.137)	(0.082)	(0.082)	(0.098)	(0.085)
Girl	-0.309	0.020	-0.160	-0.157	-0.140	-0.142	0.031	0.021
	(0.198)	(0.126)	(0.120)	(0.118)	(0.111)	(0.111)	(0.114)	(0.117)
Elected	0.198	0.123	0.259**	0.258**	0.171	0.172	0.081	0.083
	(0.131)	(0.087)	(0.100)	(0.100)	(0.102)	(0.102)	(0.097)	(0.093)
Baseline Response		0.521***		0.008		0.030		0.083*
		(0.038)		(0.034)		(0.019)		(0.043)
Constant	0.273***	0.067	0.078	0.075	0.301**	0.299**	0.047	0.041
	(0.093)	(0.067)	(0.094)	(0.094)	(0.131)	(0.130)	(0.123)	(0.126)
Observations	1,297	1,297	1,552	1,552	1,585	1,585	1,568	1,568
R-squared	0.040	0.297	0.017	0.017	0.036	0.037	0.006	0.013
Mean of dep var in control	0.104	0.141	0.036	0.041	0.223	0.258	0.078	0.080
P-value of Test of T1 + T1 * Girl	0.138	0.032	0.302	0.302	0.064	0.074	0.878	0.933
P-value of Test of T1 + T1 * Girl + T1 * Elected	0.454	0.156	0.057	0.055	0.381	0.399	0.299	0.275
P-value of Test of T2 + T2 * Girl	0.878	0.678	0.402	0.425	0.305	0.298	0.854	0.890
P-value of Test of T2 + T2 * Girl + T2 * Participant	0.556	0.249	0.932	0.916	0.046	0.043	0.287	0.304

Notes: Robust standard errors in parentheses, clustered by school.
*** p<0.01, ** p<0.05, * p<0.1
Dependent variable is the first principal component of responses to relevant questions on the endline questionnaire.

Appendix A: Boys vs. Girls (Disaggregated by Standard)

	Standard 6				Standard 7				Standard 8			
	Girls (1)	Boys (2)	Difference (3)	P-value of Test of Equality (4)	Girls (5)	Boys (6)	Difference (7)	P-value of Test of Equality (8)	Girls (9)	Boys (10)	Difference (11)	P-value of Test of Equality (12)
Scheduled Caste	0.249	0.282	-0.033	0.449	0.295	0.312	-0.017	0.683	0.218	0.384	-0.166	0.001
	(0.002)	(0.002)	(0.044)		(0.002)	(0.002)	(0.042)		(0.002)	(0.004)	(0.047)	
Scheduled Tribe	0.143	0.305	-0.162	0.000	0.112	0.268	-0.156	0.001	0.110	0.214	-0.104	0.002
	(0.001)	(0.003)	(0.039)		(0.001)	(0.003)	(0.042)		(0.001)	(0.001)	(0.031)	
Other Backwards Caste	0.457	0.300	0.157	0.002	0.426	0.278	0.148	0.002	0.451	0.276	0.175	0.002
	(0.002)	(0.002)	(0.047)		(0.002)	(0.001)	(0.044)		(0.003)	(0.001)	(0.050)	
Enrolled Previous Year	0.690	0.701	-0.011	0.862	0.913	0.894	0.020	0.475	0.937	0.924	0.012	0.575
	(0.004)	(0.002)	(0.063)		(0.001)	(0.001)	(0.027)		(0.001)	(0.001)	(0.022)	
Owns TV	0.836	0.784	0.052	0.338	0.859	0.755	0.104	0.022	0.912	0.802	0.110	0.023
	(0.002)	(0.002)	(0.053)		(0.001)	(0.002)	(0.043)		(0.001)	(0.002)	(0.046)	
Father Attended School	0.824	0.660	0.164	0.007	0.818	0.700	0.118	0.011	0.851	0.714	0.137	0.003
	(0.001)	(0.003)	(0.056)		(0.001)	(0.002)	(0.043)		(0.001)	(0.002)	(0.043)	
Mother Attended School	0.579	0.406	0.173	0.001	0.532	0.399	0.133	0.019	0.574	0.379	0.195	0.003
	(0.001)	(0.002)	(0.048)		(0.002)	(0.002)	(0.053)		(0.003)	(0.003)	(0.060)	
Education Index	-0.238	0.085	-0.322	0.008	-0.215	0.278	-0.493	0.001	-0.094	0.284	-0.377	0.026
	(0.010)	(0.011)	(0.114)		(0.015)	(0.009)	(0.128)		(0.017)	(0.009)	(0.160)	
Career Index	-0.157	0.004	-0.161	0.195	-0.183	0.104	-0.288	0.057	-0.044	0.328	-0.372	0.016
	(0.003)	(0.017)	(0.121)		(0.011)	(0.011)	(0.145)		(0.013)	(0.008)	(0.145)	
Self-confidence Index	-0.044	-0.029	-0.015	0.917	0.007	-0.049	0.055	0.619	0.016	0.101	-0.085	0.523
	(0.011)	(0.011)	(0.142)		(0.009)	(0.003)	(0.110)		(0.014)	(0.003)	(0.131)	
Gender Roles Index	0.002	-0.026	0.028	0.803	0.138	-0.158	0.296	0.014	0.210	-0.208	0.418	0.004
	(0.009)	(0.011)	(0.110)		(0.013)	(0.011)	(0.113)		(0.018)	(0.008)	(0.134)	
Number of Friends (OR)	14.741	--	--	--	16.052	--	--	--	16.686	--	--	--
	(4.102)	--			(2.071)	--			(2.497)	--		
Number of Friends (AND)	6.752	--	--	--	8.383	--	--	--	8.499	--	--	--
	(0.707)	--			(0.911)	--			(1.168)	--		

Notes: Robust standard errors in parentheses, clustered by school.

*** p<0.01, ** p<0.05, * p<0.1

Individuals *i* and *j* are "OR" friends if (at least) one names the other as a friend.

Individuals *i* and *j* are "AND" friends if they both name each other as friends.

Appendix B: Survey Attrition by Baseline Characteristics

Panel A: By Treatment Arm

Baseline Characteristic interacted with T1 and T2	Girls			Boys		
	Present for Baseline Network Survey	Present for Endline Questionnaire	Present for Endline Network Survey	Present for Baseline Network Survey	Present for Endline Questionnaire	Present for Endline Network Survey
	(1)	(2)	(3)	(1)	(2)	(3)
Standard	0.581	0.270	0.586	0.471	0.267	0.730
Age	0.097	0.148	0.077	0.412	0.931	0.681
Scheduled Caste	0.031	0.678	0.320	0.373	0.048	0.474
Scheduled Tribe	0.523	0.372	0.840	0.583	0.061	0.061
Other Backwards Caste	0.002	0.541	0.834	0.563	0.019	0.012
Enrolled Previous Year	0.023	0.059	0.048	0.983	0.314	0.365
Owns TV	0.222	0.595	0.074	0.700	0.062	0.471
Father Attended School	0.813	0.523	0.424	0.574	0.733	0.885
Mother Attended School	0.612	0.174	0.472	0.001	0.232	0.628
Education Index	0.049	0.270	0.012	0.075	0.231	0.177
Career Index	0.200	0.095	0.468	0.596	0.190	0.097
Self-confidence Index	0.031	0.429	0.450	0.378	0.013	0.057
Gender Roles Index	0.142	0.354	0.882	0.017	0.593	0.033

Panel B: Within T2

Baseline Characteristic interacted Indicator for Randomly Selected to Participate	Present for Baseline Network Survey	Present for Endline Questionnaire	Present for Endline Network Survey
	(1)	(2)	(3)
Standard	0.511	0.219	0.793
Age	0.156	0.066	0.346
Scheduled Caste	0.664	0.629	0.932
Scheduled Tribe	0.826	0.456	0.327
Other Backwards Caste	0.226	0.920	0.929
Enrolled Previous Year	0.585	0.468	0.159
Owns TV	0.062	0.910	0.155
Father Attended School	0.294	0.542	0.374
Mother Attended School	0.141	0.035	0.880
Education Index	0.405	0.510	0.090
Career Index	0.579	0.054	0.348
Self-confidence Index	0.232	0.209	0.547
Gender Roles Index	0.801	0.695	0.901

Notes: Robust standard errors in parentheses, clustered by school.
*** p<0.01, ** p<0.05, * p<0.1
Sample for Panel A is all students who were present for the Baseline Questionnaire. N = 1053 girls and 894 boys.
Sample for Panel B is all girls present for the Baseline Questionnaire in T2 schools. N = 338.
This table presents P-values of the test that the two treatment dummies interacted with baseline characteristics are jointly significant.

Appendix C: Disaggregated Baseline Network Links and Election Results

Network Definition	AND		OUT	
	(1)	(2)	(3)	(4)
Elected (OUT)	0.075**	0.045	0.065	0.036
	(0.035)	(0.028)	(0.038)	(0.033)
Elected (IN)	0.075**	0.045	0.089**	0.060**
	(0.035)	(0.028)	(0.037)	(0.029)
Elected (AND)	0.035	0.026	0.016	0.014
	(0.043)	(0.046)	(0.033)	(0.034)
T1	0.148*	0.091	0.159**	0.096
	(0.081)	(0.076)	(0.072)	(0.064)
T1 * Elected (OUT)	-0.030	-0.031	-0.092	-0.086
	(0.064)	(0.054)	(0.078)	(0.070)
T1 * Elected (IN)	-0.030	-0.031	-0.021	-0.015
	(0.064)	(0.054)	(0.054)	(0.049)
T1 * Elected (AND)	0.116	0.117	0.088	0.084
	(0.120)	(0.114)	(0.082)	(0.079)
T2	0.174**	0.097	0.161**	0.084
	(0.067)	(0.068)	(0.059)	(0.050)
T2 * Participant (OUT)	0.024	0.022	0.009	0.009
	(0.036)	(0.034)	(0.028)	(0.027)
T2 * Participant (IN)	0.024	0.022	-0.001	-0.000
	(0.036)	(0.034)	(0.024)	(0.021)
T2 * Participant (AND)	0.064	0.060	0.054*	0.051
	(0.047)	(0.049)	(0.028)	(0.030)
Constant	0.257***	1.464*	0.465***	1.368**
	(0.061)	(0.843)	(0.063)	(0.640)
School Size Controls Included	NO	YES	NO	YES
Observations	27,418	27,418	27,418	27,418
R-squared	0.055	0.073	0.042	0.055
Mean of Dep var in Control	0.286	0.286	0.493	0.493
P-value of Test of T1 + T1* Elected (OUT)	0.278	0.568	0.564	0.931
P-value of Test of T1 + T1* Elected (IN)	0.278	0.568	0.033	0.199
P-value of Test of T1 + T1* Elected (OUT) + T1*Elected (IN) + T1*Elected (AND)	0.029	0.154	0.064	0.323
P-value of Test of T2 + T2*Participant (OUT)	0.005	0.055	0.010	0.072
P-value of Test of T2 + T2*Participant (IN)	0.005	0.055	0.016	0.162
P-value of Test of T2 + T2*Participant (OUT) + T2*Participant(IN) + T2*Participant(AND)	0.000	0.001	0.001	0.007

Notes: Robust standard errors in parentheses, clustered by school.

*** p<0.01, ** p<0.05, * p<0.1

Dependent variable is existence of friendship at Baseline, under appropriate definition.

School Size Controls include linear, quadratic, and cubic in number of students enrolled at beginning of school year.

Appendix D: Disaggregated Program ITT Effects on Network Formation

Panel A: Not Friends at Baseline

Network Definition	OR			AND		
	(1)	(2)	(3)	(4)	(5)	(6)
Elected (OR)	0.107**	0.081*	0.048	0.073**	0.065**	0.044
	(0.042)	(0.043)	(0.047)	(0.029)	(0.029)	(0.029)
Elected (AND)	0.217***	0.135**	0.087	0.119**	0.069	0.043
	(0.069)	(0.057)	(0.069)	(0.049)	(0.051)	(0.057)
T1	0.126	0.104	0.002	0.075*	0.060**	-0.019
	(0.082)	(0.070)	(0.093)	(0.037)	(0.028)	(0.053)
T1 * Elected (OR)	-0.104	-0.109	-0.087	-0.114**	-0.109**	-0.096*
	(0.083)	(0.093)	(0.097)	(0.045)	(0.048)	(0.047)
T1 * Elected (AND)	-0.040	0.021	0.056	-0.071	-0.039	-0.011
	(0.118)	(0.110)	(0.112)	(0.102)	(0.111)	(0.110)
T2	0.229***	0.222***	0.130*	0.059	0.044	-0.010
	(0.063)	(0.059)	(0.065)	(0.045)	(0.043)	(0.041)
T2 * P (OR)	-0.064*	-0.088**	-0.091**	0.050	0.042	0.039
	(0.037)	(0.043)	(0.044)	(0.042)	(0.034)	(0.032)
T2 * P (AND)	-0.061	-0.067	-0.067	0.134	0.146	0.141
	(0.090)	(0.098)	(0.102)	(0.108)	(0.093)	(0.093)
Constant	0.388***	0.316***	2.128**	0.085***	0.065***	1.468**
	(0.046)	(0.047)	(0.898)	(0.023)	(0.020)	(0.571)
Baseline Network Controls Included	NO	YES	YES	NO	YES	YES
School Size Controls Included	NO	NO	YES	NO	NO	YES
Observations	3,140	3,140	3,140	3,140	3,140	3,140
R-squared	0.048	0.091	0.102	0.033	0.088	0.097
Mean of Dep var in Control	0.631	0.631	0.631	0.325	0.325	0.325
P-value of Test of T1 + T1 * Elected (OR)	0.837	0.964	0.450	0.414	0.301	0.046
P-value of Test of T1 + T1 * Elected (OR) + T1 * Elected (AND)	0.811	0.821	0.728	0.231	0.377	0.236
P-value of Test of T2 + T2 * Participant (OR)	0.010	0.060	0.602	0.062	0.149	0.622
P-value of Test of T2 + T2 * Participant (OR) + T2 * Participant (AND)	0.350	0.602	0.829	0.013	0.009	0.048

Panel B: "OR" Friends but not "AND" Friends at Baseline

Network Definition	OR			AND		
	(1)	(2)	(3)	(4)	(5)	(6)
Elected (OR)	0.090**	0.066*	0.050	0.072	0.057	0.046
	(0.038)	(0.035)	(0.031)	(0.044)	(0.043)	(0.039)
Elected (AND)	0.100*	0.053	0.021	0.178***	0.131**	0.106**
	(0.050)	(0.043)	(0.042)	(0.055)	(0.051)	(0.051)
T1	0.042	0.030	-0.015	-0.005	-0.013	-0.041
	(0.091)	(0.077)	(0.087)	(0.059)	(0.054)	(0.083)
T1 * Elected (OR)	0.069	0.075	0.076	0.072	0.083	0.082
	(0.054)	(0.053)	(0.048)	(0.059)	(0.062)	(0.052)
T1 * Elected (AND)	-0.058	-0.081	-0.068	-0.065	-0.067	-0.059
	(0.113)	(0.116)	(0.116)	(0.158)	(0.163)	(0.167)
T2	0.171**	0.193**	0.132*	0.125	0.138	0.091
	(0.073)	(0.075)	(0.069)	(0.084)	(0.087)	(0.095)
T2 * P (OR)	-0.078**	-0.095**	-0.086**	-0.088	-0.091	-0.083
	(0.037)	(0.036)	(0.036)	(0.061)	(0.057)	(0.060)
T2 * P (AND)	0.034	-0.002	0.004	0.164**	0.132***	0.137***
	(0.062)	(0.058)	(0.061)	(0.062)	(0.044)	(0.041)
Constant	0.588***	0.481***	1.746*	0.255***	0.182***	1.167
	(0.062)	(0.068)	(1.001)	(0.048)	(0.048)	(1.112)
Baseline Network Controls Included	NO	YES	YES	NO	YES	YES
School Size Controls Included	NO	NO	YES	NO	NO	YES
Observations	6,152	6,152	6,152	6,152	6,152	6,152
R-squared	0.042	0.084	0.093	0.038	0.067	0.072
Mean of Dep var in Control	0.631	0.631	0.631	0.325	0.325	0.325
P-value of Test of T1 + T1 * Elected (OR)	0.119	0.097	0.442	0.354	0.314	0.678
P-value of Test of T1 + T1 * Elected (OR) + T1 * Elected (AND)	0.648	0.847	0.958	0.988	0.987	0.918
P-value of Test of T2 + T2 * Participant (OR)	0.131	0.094	0.408	0.591	0.460	0.903
P-value of Test of T2 + T2 * Participant (OR) + T2 * Participant (AND)	0.181	0.306	0.604	0.022	0.010	0.054

Panel C: "AND" Friends at Baseline

Network Definition	OR			AND		
	(1)	(2)	(3)	(4)	(5)	(6)
Elected (OR)	0.038	0.034	0.035	0.124***	0.107***	0.107***
	(0.032)	(0.029)	(0.029)	(0.036)	(0.029)	(0.031)
Elected (AND)	0.082**	0.079	0.081	0.106**	0.092*	0.095*
	(0.039)	(0.049)	(0.050)	(0.041)	(0.047)	(0.047)
T1	-0.014	-0.005	-0.007	-0.103*	-0.099**	-0.110*
	(0.051)	(0.045)	(0.052)	(0.053)	(0.043)	(0.057)
T1 * Elected (OR)	0.028	0.022	0.027	0.000	-0.012	0.001
	(0.051)	(0.044)	(0.045)	(0.058)	(0.049)	(0.048)
T1 * Elected (AND)	-0.010	-0.015	-0.016	0.038	0.031	0.027
	(0.048)	(0.053)	(0.055)	(0.061)	(0.067)	(0.070)
T2	0.048	0.053	0.061	0.063	0.060	0.073
	(0.055)	(0.050)	(0.054)	(0.085)	(0.083)	(0.089)
T2 * P (OR)	-0.038	-0.034	-0.036	-0.071	-0.076	-0.080*
	(0.028)	(0.028)	(0.028)	(0.046)	(0.047)	(0.044)
T2 * P (AND)	0.054	0.059**	0.059**	0.042	0.041	0.044
	(0.033)	(0.027)	(0.029)	(0.050)	(0.041)	(0.042)
Constant	0.821***	0.791***	0.781**	0.524***	0.515***	0.566
	(0.032)	(0.042)	(0.349)	(0.048)	(0.072)	(0.616)
Baseline Network Controls Included	NO	YES	YES	NO	YES	YES
School Size Controls Included	NO	NO	YES	NO	NO	YES
Observations	6,286	6,286	6,286	6,286	6,286	6,286
R-squared	0.020	0.048	0.049	0.042	0.087	0.090
Mean of Dep var in Control	0.631	0.631	0.631	0.325	0.325	0.325
P-value of Test of T1 + T1 * Elected (OR)	0.708	0.648	0.658	0.084	0.042	0.097
P-value of Test of T1 + T1 * Elected (OR) + T1 * Elected (AND)	0.927	0.950	0.931	0.463	0.333	0.404
P-value of Test of T2 + T2 * Participant (OR)	0.827	0.651	0.609	0.904	0.798	0.921
P-value of Test of T2 + T2 * Participant (OR) + T2 * Participant (AND)	0.066	0.021	0.031	0.508	0.584	0.493

Notes: Robust standard errors in parentheses, clustered by school.

*** p<0.01, ** p<0.05, * p<0.1

Dependent variable is existence of friendship at Endline.

Baseline Network Controls include answer to all baseline network survey questions, as well as interactions between (IN) and (OUT).

School Size Controls include linear, quadratic, and cubic in number of students enrolled in the school at beginning of school year.

Appendix E: Program Effects on Friend Count (Degree)

Network Definition	OR			AND		
	(1)	(2)	(3)	(4)	(5)	(6)
Elected	-0.782	-0.307	2.050	2.324***	2.173***	3.050***
	(1.483)	(1.051)	(1.329)	(0.833)	(0.772)	(1.041)
T1	-4.408	-3.120	3.908	-1.555	-1.421	1.373
	(5.753)	(4.299)	(2.929)	(1.802)	(1.562)	(1.496)
T1 * Elected	0.734	1.226	-0.695	-0.299	-0.134	-0.979
	(2.304)	(2.033)	(2.073)	(1.381)	(1.388)	(1.477)
T2	-0.806	0.138	5.646*	1.766	1.502	3.629*
	(6.041)	(4.716)	(3.169)	(2.747)	(2.233)	(1.798)
T2 * Participant	-1.618	-1.265	-1.643	-0.689	-0.790	-0.954
	(1.623)	(1.295)	(1.357)	(1.266)	(1.011)	(1.004)
Baseline Friend Count (OR)		0.621***	0.490***		0.171***	0.125**
		(0.083)	(0.074)		(0.038)	(0.050)
Baseline Friend Count (AND)		-0.262*	-0.053		0.178*	0.251*
		(0.150)	(0.183)		(0.103)	(0.132)
Constant	23.721***	15.100***	-58.259**	9.678***	5.675***	-20.191
	(5.164)	(3.821)	(27.182)	(1.527)	(1.308)	(14.733)
School Size Controls Included	NO	NO	YES	NO	NO	YES
Observations	1,376	1,376	1,376	1,376	1,376	1,376
R-squared	0.010	0.167	0.245	0.027	0.177	0.218
Mean of Dep var in Control	21.676	21.676	21.676	9.408	9.408	9.408
P-value of Test of T1 + T1 * Elected	0.528	0.685	0.362	0.494	0.519	0.856
P-value of Test of T2 + T2 * Participant	0.671	0.805	0.207	0.633	0.728	0.086

Notes: Robust standard errors in parentheses, clustered by school.

*** p<0.01, ** p<0.05, * p<0.1

Dependent variable is number of friends at Endline under appropriate definition.

School Size Controls include linear, quadratic, and cubic in number of students enrolled in the school at beginning of school year.