Essays on Social Networks in Development Economics

by

Michael Alan Griffith

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in the University of Michigan
2017

Doctoral Committee:

Professor Jeffrey A. Smith, Chair
Professor Daniel A. Ackerberg
Associate Professor Tanya Rosenblat
Associate Professor Rebecca L. Thornton, University of Illinois

Michael Alan Griffith

alangrif@umich.edu

ORCID iD: 0000-0001-5537-768X

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

LIST OF APPENDICES

Appendix

ABSTRACT

Over the past decades, economists have increasingly taken note of the importance of social networks in determining choices and behaviors in a wide variety of settings. More recently, researchers have taken note of the role of network endogeneity–that is, people choosing their peer groups–in informing our understanding of many relationships. This dissertation consists of three chapters that investigate the role of social networks in economics, with a focus on applications in developing-country settings. Chapter 1 derives a new methodology to predict the effects of policies that assign people to groups, while taking account of the fact that actors choose their peers and peers affect outcomes. Chapter 2 investigates the sensitivity of estimate of the effect of peers to a particular and widely-used data collection process. In joint work, Chapter 3 investigates the effects of a particular development program targeted at adolescent girls, taking account of the fact that program assignment affects patterns of interaction. Taken together, these chapters form the beginning of a long-term research agenda that seeks to understand the causes and effects of social networks in development economics.

CHAPTER 1

Random Assignment with Non-Random Peers: A Structural Approach to Counterfactual Treatment Assessment

Abstract

Recent efforts by economists to leverage peer effects by creative peer assignment have come up short due in part to endogenous peer selection. That is, even conditional on random assignment, agents choose their peers, and failure to account for this selection may crucially bias estimates of peer effects and, in turn, predictions of the effects of alternative policies. To address this shortcoming, I build a two-part model in which (1) agents form a network; (2) conditional on the realized network, outcomes are determined by a process that allows for non-linear peer effects. To overcome difficulties in identification and estimation of network-formation games, agents in my model make continuous linking decisions subject to a budget constraint. I show that, under certain conditions, this model has a unique strictly positive equilibrium, which can then be used for identification and estimation. In modeling peer effects, I explicitly model network endogeneity as an omitted variable problem, and further propose a method to recover these omitted variables in estimating the network-formation game. I estimate the parameters of the two-part model using innovative data on networks and outcomes from a randomized study in Rajasthan, India, then show that the model performs well in matching predictions to realized out-of-sample outcomes. This paper makes important contributions to the methodology of peer effects estimation as well as the theory and econometrics of network formation, while providing an important link between structural and experimental approaches to policy evaluation.

1.1 Introduction

A rich literature in economics and related fields shows that individual choices and outcomes are not independent of the choices, outcomes, and characteristics of those they interact with. Further, a growing number of papers exploiting random assignments credibly makes the case that these peer effects can be given a causal interpretation (Epple and Romano, 2011; Sacerdote, 2011). For example, prominent studies have exploited random assignment to dorms (Sacerdote, 2001; Stinebrickner and Stinebrickner, 2006), university class sections (DeGiorgi, Pelllizzari and Radaelli, 2010), second-grade classrooms in rural Kenya (Duflo, Dupas and Kremer, 2011), and squadrons at the Air Force Academy (Carrell, Fullerton and West, 2009; Carrell, Hoekstra and West, 2011). These studies tend to find large and statistically significant peer effects on a variety of outcomes (Epple and Romano, 2011). This robust evidence for the existence of peer effects suggests that creative peer assignment may be a powerful policy tool to influence individual choices and outcomes. That is, if peer effects are sufficiently strong, simply changing the composition of peer groups may substantially change measurable outcomes.

However, efforts to leverage peer effects to improve outcomes have proven difficult due to, among other reasons, a failure to account for endogenous sorting (see, e.g., Angrist, 2014; Carrell, Sacerdote and West, 2013). That is, policy interventions designed to change peer group composition may also affect patterns of interaction: even with random assignment, agents still may choose with whom they interact. The importance of this channel is highlighted by a number of recent studies documenting experimental interventions that change network structure (see Comola and Prina, 2014; Delavallade, Griffith and Thornton, 2016; Vasilaky and Leonard, 2014). Accordingly, in such settings, efforts to predict the effects of alternative assignments must carefully account for the effects of said assignments not only *through* peers but also on the *choice* of peers.[1] Thus, several recent papers have suggested

[1] Of course, if the goal is simply to learn the effect of alternative policies on some outcome, researchers may design studies to cover a wide range of possible alternative assignments while remaining agnostic about network endogeneity, but such a strategy is often impractical due to institutional or funding limitations.

pairing models of peer effects with models of peer choice that can be taken to data (Blume et al., 2015; Graham, 2014*b*).

Network formation models are notoriously difficult to estimate, however, due to related issues of theory, identification, and computation.[2] The bulk of the theory on network formation posits links as binary decisions: they either exist or do not (but see Bloch and Dutta, 2009; Baumann, 2016). The discrete nature of these games necessitates specifying alternatives to Nash equilibrium such as pairwise stability. These games tend to be characterized by the existence of multiple equilibria, a feature that complicates identification, leading to partial identification (see de Paula, Richards-Shubik and Tamer, 2016; Sheng, 2012) or the need to specify complex equilibrium selection rules (see Badev, 2013; Christakis et al., 2010; Mele, 2010). Finally, the discrete nature of the problem implies the need to calculate high-dimensional inequalities (see, e.g. Sheng, 2012), leading to a curse of dimensionality in estimation.

To surmount these difficulties, I model the network formation process as a static, simultaneous move game in which players make *continuous* linking decisions. Additionally, linking decisions are made subject to a budget constraint, which necessarily builds in tradeoffs between forming links. Importantly, I show that the model has a unique strictly positive Nash equilibrium, in which all agents link positively to all other agents. This crucial feature facilitates identification and estimation of the network formation game without reference to partial identification or complicated equilibrium-selection procedures. The strictly positive equilibrium is characterized by linear best-response functions which can then be used for identification and estimation. Further, the tradeoffs implied by the budget constraint motivate the relevance of a budget-set instrument to identify parameters of the network formation game. With sufficient variation in exogenous characteristics, parameters of the network formation model are point identified. Most importantly, individual-specific unob-

One notable exception is Booij, Leuven and Oosterbeek (2016).

[2] Chandrasekhar (2015) and Graham (2014*b*) provide comprehensive overviews of the current literature on the identification and estimation of network formation models.

served variables are identified as the size of each observed network grows.

Next, I model outcomes conditional on the realized network. I generalize a reduced-form version of the linear-in-means model (see Manski, 1993) by including additively-separable unobserved or "latent" effects (Goldsmith-Pinkham and Imbens, 2013; Hsieh and Lee, 2016; Jackson, 2014) as well as non-linear peer effects (Carrell, Sacerdote and West, 2013). This approach explicitly models network endogeneity as an omitted-variable problem, and failure to account for the latent characteristics biases estimates of the peer effects model. Crucially, the unobserved variables that cause bias due to network endogeneity are the same unobserved individual-specific parameters that are identified in the network-formation process. Conditional on their identification, I show that the parameters of the peer effects model are identified even in the presence of network endogeneity.

Next, I take these identification results to data gathered as part of a randomized trial of a girls' empowerment program in rural Rajasthan, India. The study design consists of a treatment that randomly assigns 13 girls (out of approximately 44) in each of 10 schools to participate in an after-school program, while control schools do not receive any programming. As part of this effort, we collected especially rich data on social networks, consisting of a pairwise network census with detailed data on network connections along a number of dimensions. From this rich data, I construct a continuous measure of connectedness.

I first estimate the network-formation model to recover the individual-specific unobservables. These estimates indicate that these unobservables are quite important in determining network structure. I then plug these into the peer-effects model, which allows for consistent estimation of the parameters of that model. For two program outcomes, I show that the individual-specific unobservables significantly affect realized outcomes, and statistical tests strongly reject simpler models that do not account for them. That is, network *exogeneity* is a nested special case of my model, and this case can be rejected with a high degree of certainty.

As a further check, I employ the estimated parameters in simulating outcomes to compare to realized out-of-sample means. This step shows that the model performs well in out-of-sample prediction, a validation step proposed by Todd and Wolpin (2006). Finally, I simulate predicted outcomes under three different assignment rules, revealing that the assignment rule substantially affects one program outcome of interest but has little effect on the other.

This paper's primary contribution lies in providing a method for estimating peer effects models in the presence of endogenous network formation. Such estimates can then be used to predict the effects of alternative policies while simultaneously accounting for the effect of those policies on network structure. As a necessary step in developing this method, I make two additional contributions. First, I advance the theoretical literature on network formation, particularly in the context of agents making continuous linking decisions. Second, building upon this novel theoretical model, I provide an important advance in the econometric literature on the identification of network-formation games. Finally, in the empirical application, I contribute to the literature building bridges between structural and experimental approaches to program evaluation, especially in development contexts (see, e.g., Attanasio, Meghir and Santiago, 2011; Duflo, Hanna and Ryan, 2012; Todd and Wolpin, 2006), as well as adding to work comparing randomized to non-randomized assignments, as demonstrated, for example, in Shadish, Clark and Steiner (2008).

This paper proceeds as follows. Section 1.2 describes the program under study as well as deriving a number of key reduced-form facts to motivate features of the structural model. Section 1.3 provides the peer effects model that posits network endogeneity as an omitted-variable problem. Section 1.4 then develops the network-formation model as a means of controlling for these structural unobservables. Applying the identification results, Section 1.5 presents results of structural estimation of the network formation and peer effects models. Using these parameter estimates, Section 1.6 compares predicted outcomes to those of a realized out-of-sample group of schools. Section 1.7 then simulates outcomes under alternative assignment regimes whereby students are assigned preferentially based upon either predicted outcomes or estimated unobservables. Section 1.8 concludes.

1.2 Empirical Setting and Data Description

This section describes the empirical setting and the experimental design. Through this exercise, we learn that the program has negative but insignificant effects on outcomes but strongly significant and substantively meaningful effects on network structure. I describe how I construct the continuous network measure, then show that the constructed continuous measure of networks is consistent with binary measures that are typically used in peer-effects models and the economic analysis of networks. Finally, I use the network data to demonstrate stylized facts that motivate features of the network formation model developed in Section 1.4.

1.2.1 Girls' Empowerment Program

The NGO Educate Girls operates the Bal Sabha program in government schookls in rural Rajasthan, India. As part of the program, 13 girls in grades 6 through 8 are chosen to form a Bal Sabha (Girls' Parliament). The program focuses on developing so-called "soft skills." These skills include leadership and self-confidence as well as attitudes and aspirations about education, age at marriage, and gender roles. The larger goal of the program is for girls to employ these skills as a means of overcoming barriers to their own education, such as early marriage.

The intervention consists of a series of five "games" during which village volunteers work through increasingly difficult scenarios. Through activities such as role playing, girls are taught to develop their own voice in difficult situations such as, for example, their parents desiring to have them marry young or end their schooling. The parliaments meet biweekly over a span of approximately six months during the academic year. The total intervention time is approximately 25-50 hours. Girls chosen to participate were encouraged to share their learning and experiences with their classmates who were not participants, and effects spilling over to non-participants is a key feature of the implementing organization's theory of change.

Under the NGO's preferred assignment rule, the 13 participants girls in each school are chosen through elections involving all students in grades 6-8, including boys. These elections lead to non-random selection, a fact that is documented in Delavallade, Griffith and Thornton (2016).

1.2.2 Study Design and Data Collection

As part of the rollout of the program to a new district during the 2013-14 academic year, a study team designed and implemented a randomized trial. A sample of thirty schools was chosen in two administrative blocks. None of these schools had ever had a Bal Sabha prior to the start of the study.

Prior to treatment assignment, three data collection activities in each school were conducted in each school. First, Bal Sabha elections were held in all schools, including those that would later serve as controls. Second, girls in each school filled out an extensive questionnaire that gathered background demographic information as well as data on attitudes, aspirations, and expectations along a number of dimensions. Third, prior to treatment assignment, a pairwise network census was also collected among all girls in each of the 30 schools, the form of which is described below.

After baseline data collection, schools were assigned to one of three treatment groups. In Random Treatment schools, girls were randomly chosen to participate. In Elected Treatment schools, the program was conducted with the girls chosen by election, as is customary for the program as implemented by Educate Girls. Finally, Control schools did not receive the program in any form.

The program was implemented over a period of approximately six months. During this time, village volunteers trained by the NGO led the 13 participating girls through the games-based curriculum. At the conclusion of program implementation, enumerators returned to each school to conduct an endline survey that measured outcomes similar to those measured at baseline. Further, in order to assess the effect of the program on networks, we conducted an additional pairwise network census. Accordingly, this data allows us to measure the program's effects on both endline outcomes—as measured by aspirations and attitudes—and endline networks.

1.2.3 Demographics and Outcomes at Baseline

Table 1.1 provides descriptive statistics for the full sample of 1319 girls at baseline.[3] Note that approximately 28 percent of the girls were elected to participate, out of an average of approximately 44 girls in each school. Enrollment is slightly skewed toward girls in Grade 6 (the omitted category). Finally, note substantial variation in caste, as 37% of the sample are members of Scheduled Castes/Scheduled Tribes, while 44.5% are in Other Backwards Castes. The omitted caste category, General or upper castes, comprises 18.5% of the sample. From these means, we see that there are fewer girls in higher grades, suggestive of school dropout, as well as large lower-caste populations.

This paper focuses on two outcomes: educational aspirations and attitudes about gender roles. These outcome measures are constructed as the normalized first principal component of all relevant survey questions.[4] Girls have higher educational aspirations if, for example, they indicate they would like to complete university, as compared to stating they would like to complete only eighth grade. Girls have higher Gender Roles attitudes if, for example, they say it is okay for a wife to disagree with her husband in public.

Baseline outcomes are summarized in Table 1.1, Panel B. Since the mean of the outcome variables is zero by construction in the data among *all* students (including boys), these means indicate that girls have below average Educational Aspirations and above average Gender Roles attitudes. This conforms to our priors that girls have lower Educational Aspirations than boys but higher Gender Roles attitudes.[5]

[3] The sample consists of all girls who have non-missing data on the covariates in Panel A. This consists of more than 99% of eligible girls.

[4] That is, among all students in the sample, the mean is set to zero with variance of one.

[5] Baseline balance is presented in Appendix B. Table B.1 shows balance across treatment arms, while Table B.2 shows within-school balance between girls (randomly) selected to participate in Random Treatment and those not selected.

Table 1.1: Baseline Variable Descriptives

	Mean	S.D.
Panel A: Baseline Covariates		
Elected	0.281	0.449
Grade 7	0.318	0.466
Grade 8	0.306	0.461
Scheduled Caste	0.252	0.435
Scheduled Tribe	0.118	0.323
Other Backward Caste	0.445	0.497
Panel B: Baseline Outcomes		
Educational Aspirations	-0.197	1.036
Gender Roles	0.115	0.984

Notes: Robust standard errors in parentheses, clustered by school. Sample is all girls in all schools. N = 1319 in 30 schools.

Table 1.2 presents regression results of baseline outcomes on the covariates in Panel A of Table 1.1. The results show that these outcomes vary substantially by baseline characteristics. For example, elected girls have 0.168 standard deviations higher Educational Aspirations on average, while lower caste girls (SC, ST, and OBC) have substantially lower levels of both baseline outcomes. Accordingly, if the program is to be targeted at those most "at need," it may make sense to target lower caste girls for participation in the program rather than the more popular girls who are chosen by election.

1.2.4 Baseline Network Descriptives, and a Continuous Measure of Connectedness

Here, I describe the network data. Through this, we see that the data consists of a large number of binary network measures, from which I construct a continuous measure of connectedness. Further, I show that patterns of links are qualitatively similar for the continuous measure and for the binary measure that is often used in the literature. This in turn suggests the reasonableness of the constructed continuous measure that is necessary for the structural model developed in later sections.

Table 1.2: Baseline Outcome Heterogeneity

	Educational Aspirations (1)	Gender Roles (2)
Elected	0.168**	-0.042
	(0.078)	(0.073)
Grade 7	0.024	0.125*
	(0.080)	(0.070)
Grade 8	0.077	0.172**
	(0.099)	(0.084)
Scheduled Caste	-0.234*	-0.337***
	(0.118)	(0.120)
Scheduled Tribe	-0.185	-0.544***
	(0.118)	(0.111)
Other Backwards Caste	-0.284***	-0.078
	(0.087)	(0.102)
Constant	-0.068	0.219**
	(0.092)	(0.080)
R-squared	0.018	0.040

Notes: N = 1,319 in 30 schools in all specifications. Robust standard errors in parentheses, clustered by school. *** p<0.01, ** p<0.05, * p<0.1.

The network data come from a pairwise network census. This procedure consisted of each girl in the sample answering a series of binary questions about every other girl in her school in grades 6 through 8. I categorize these variables by whether they are choices, such as being friends, or static variables, such as living in close proximity. These are described in Table 1.3 in Panels A and B, respectively. I have highlighted the "She is a friend" measure in gray, as that is the link definition commonly reported in the literature. Note that the friendship networks are quite dense: on average, girls indicate that 45.8 percent of other girls in their school are friends, as shown in the shaded row of Table 1.3. Other measures of connectedness, on the other hand, suggest less dense networks. For example, only 23.8% of girls say that they have spent time outside school with the other in the past week.

Table 1.3: Endline Network Variable Descriptives

	Mean	In/Out Correlation	Factor Loading
Panel A: Choice Network Variables			
She is a friend	0.458	0.334	0.324
I speak with her regularly	0.373	0.291	0.343
In the past week, spent time outside school	0.238	0.244	0.310
I think she is clever	0.397	0.176	0.334
She has a lot of friends / is popular	0.384	0.257	0.347
She is very shy/quiet	0.399	0.139	0.270
I think she is very confident	0.275	0.214	0.367
I wish I could be like her	0.210	0.171	0.337
I can trust her to keep my secrets	0.245	0.244	0.359
Panel B: Static Network Variables			
She is a relative	0.154	0.434	
We are in the same caste	0.208	0.704	
I can walk to her home in less than 10 minutes	0.248	0.282	

Notes: Sample is all pairs of students. N = 78,238 in 30 schools. Missing data imputed via iterative EM algorithm (see Appendix D). First principal component explains 47.3% of variation.

While the bulk of the economics literature on networks treats links as binary, the additional measures of connectedness allow me to capture more variation in link intensity.[6] In

[6] If connectedness is indeed a latent continuous measure, an additional motivation for use of the first

order to exploit this additional information—and as an input into the structural estimation process described in later sections of this paper—I construct a "continuous" measure of connectedness by collapsing the measures in Table 1.3, Panel A into a single index. While this provides for more exploitable variation, these variables are highly correlated with each other. To account for the covariance structure, I take the first principal component, scaled such that the constructed continuous measure has minimum zero and unit variance. The final column of Table 1.3 provides the factor loadings for each variable in Panel A, and the first component accounts for 47.3% of the variation in the included variables.

In Table 1.4, I compare the continuous measure of connectedness to the binary one. For the latter, I follow the bulk of the literature in using the student's response to "She is a friend" as a binary link measure. Panel A shows the probability that a student in the group identified on the y-axis indicates an individual on the x-axis is a friend. For example, the likelihood that a member of a Scheduled Caste names another Scheduled Caste member as a friend is 57.7%, while the likelihood of her naming a member of General Castes is 42.2%. Comparison of (shaded) elements along the diagonal with others in the same row suggests individuals are much more likely to claim as friends others of their own population grouping. The final column provides the p-value of a test of the equality between the probability of an individual in that row indicating a same-type other individual is a friend with the probability of her indicating an individual in a different category is a friend. Note that this test suggests strong homophily among members of Scheduled Castes and General Castes, but provides weaker evidence for Scheduled Tribes and Other Backwards Castes under the binary link definition.

Panel B performs the same exercise as Panel A, except with means of the continuous measure of connectedness. Therefore, in Panel B the means in the table represent the mean connectedness value that an individual in the group on the y-axis assigns to an individual on the x-axis. From this, we see similar patterns of homophily: Scheduled Castes and General

principal component is to reduce measurement error (see Black and Smith, 2006).

Table 1.4: Homophily by Population Group

Panel A: Binary Network Definition	SC	ST	OBC	General	P-value of Test for Homophily
SC	0.577	0.386	0.407	0.422	0.000
	(0.021)	(0.029)	(0.026)	(0.035)	
ST	0.407	0.484	0.412	0.348	0.131
	(0.039)	(0.039)	(0.066)	(0.041)	
OBC	0.388	0.374	0.440	0.446	0.068
	(0.045)	(0.067)	(0.055)	(0.038)	
General	0.362	0.314	0.402	0.564	0.000
	(0.034)	(0.040)	(0.037)	(0.025)	

Panel B: Continuous Link Intensity Definition	SC	ST	OBC	General	P-value of Test for Homophily
SC	1.294	0.864	0.895	0.936	0.003
	(0.087)	(0.078)	(0.102)	(0.093)	
ST	0.978	1.363	1.023	0.844	0.004
	(0.098)	(0.048)	(0.185)	(0.100)	
OBC	0.883	0.848	1.000	0.987	0.117
	(0.110)	(0.151)	(0.138)	(0.086)	
General	0.822	0.710	0.915	1.399	0.000
	(0.083)	(0.078)	(0.095)	(0.111)	

Notes: Observations at the link level: N = 78,238 in 30 schools. Values indicate mean value of link for individual in group on the y-axis with respect to individual in group on the x-axis. Robust standard errors in parentheses, clustered by school. Final column presents p-value of test that mean value of link is equal for same type and other types. For example, for SC, it tests that the mean link value to other SCs is the same as the average link value of ST, OBC, and General pooled together. That is, it tests the equality of the mean on the diagonal to the pooled mean of off-diagonal elements within the same row. SC = Scheduled Caste, ST = Scheduled Tribe, OBC = Other Backwards Caste.

castes continue to show substantial homophily, with weaker evidence for Other Backwards Castes. Interestingly, the continuous measure also suggests that we can reject the null of no homophily for Scheduled Tribes, in contrast to the case of the binary link measure. In all, these results suggest that the continuous measure of connectedness reflects similar network patterns to the binary "She is a friend" measure that has received the bulk of attention in the literature on the economics of networks.

1.2.5 Stylized Facts about Networks

Table 1.5 presents additional facts about networks. The regression results are presented for descriptive purposes, making no claims as to causation, in order to motivate features of the network formation model in the following sections.

First, Panel A shows that the size of an individual's network is increasing in the size of her school, as defined by the number of girls in grades 6-8. For Column (1), with the binary link definition, this is measured by the simply the number of friends she claims to have. A school having one additional girl is associated with other students having, on average 0.576 more friends. In Column (2), with the continuous link definition, the dependent variable is the sum of her scalar link measures with reference to all other girls in her school. The coefficient on School Size indicates that an additional student is associated with 0.392 higher sum of links.

Second, Panel B shows that average link value is decreasing in school size. This is indicated by the negative and highly significant coefficient on the School Size variable. This suggests that there may be tradeoffs in linking strategies, and that each individual's many linking decisions are not independent of each other. This casts doubt on models of link formation employed, for example, in Goldsmith-Pinkham and Imbens (2013) and Comola and Prina (2014), which effectively assume independence of links, implying that the coefficient on School Size in Panel B would be zero.

Third, Panel C shows that linking decisions are complementary but not symmetric. That is, individual i's decision to link to j (i's out-link with respect to j) depends upon j's decision to link to i (i's in-link with respect to j), as shown by the highly significant positive coefficients in Panel C.[7] While linking strategies are clearly complementary, neither of the two measures (binary and continuous) is symmetric, as would be indicated by a coefficient of unity.[8] That is, while in- and out-links are correlated, they are not symmetric.

1.2.6 Reduced-Form Treatment Effects

Before proceeding to structural modeling, here I present reduced-form treatment effects, on both outcomes and networks. I restrict this exercise to Random Treatment and Control schools so that we can interpret differences between those chosen for participation under Random Treatment and those not chosen as causal. These results reveal that the program has negative but insignificant effects on endline outcomes, especially for participants, and that selection to participation has significant effects on networks. The structural model developed in later sections accounts for these features of the data.

1.2.6.1 Effects on Outcomes

First, I estimate reduced-form treatment effects with specifications as in Equation (1.1). The omitted category in these regressions is all students in Control schools.

$$\begin{aligned} y_{is} = \beta_0 &+ \beta_1 RandomTreat_s \times Participant_{is} \\ &+ \beta_2 RandomTreat_s \times NonParticipant_{is} + \epsilon_{is} \end{aligned} \tag{1.1}$$

Table 1.6 shows reduced-form treatment effects. While noting possible lack of statistical power to detect small differences, first observe that the point estimates of the program's effects are negative in all specifications. Further, the point-estimated effects of approximately

[7] That is, i's in-link with respect to j is the same as j's out-link with respect to i, etc.

[8] Even if links are indeed symmetric, measurement error in the network measure would tend to attenuate the estimated coefficient away from one.

Table 1.5: Network Size and Compementarity

Panel A: Relationship between School Size and Link Count

Network Definition	Binary	Continuous
	(1)	(2)
School Size	0.576***	0.392***
	(0.030)	(0.029)
Constant	20.988***	8.448***
	(2.913)	(1.510)
R-squared	0.481	0.642

Panel B: Relationship between School Size and Link Value

Network Definition	Binary	Continuous
	(1)	(2)
School Size	-0.002***	-0.004***
	(0.000)	(0.001)
Constant	0.684***	1.289***
	(0.021)	(0.082)
R-squared	0.024	0.035

Panel C: Relationship between In- and Out-Link Values

Network Definition	Binary	Continuous
	(1)	(2)
In-Link Value	0.129***	0.224***
	(0.016)	(0.039)
Constant	0.471***	0.729***
	(0.033)	(0.042)
R-squared	0.017	0.050

Notes: N = 1,19 in 30 schools in Panel A, N = 78,238 in 30 schools in Panels B and C. Robust standard errors in parentheses, clustered by school. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Dependent variable for Panel A is sum (or count) of links under appropriate definition. Dependent variable for Panels B and C is value of out-link under appropriate definition. Unit of observation is individual student in Panel A, dyad (pair of students) in Panels B and C.

-0.2 standard deviations are substantively meaningful, at least among participants. Additionally, the effect on both outcomes is more negative for participants than non-participants, although both are insignificant in all specifications.

Table 1.6: Reduced-From Treatment Effects

	Education		Gender Roles	
	(1)	(2)	(3)	(4)
Random Treat × Participant	-0.181	-0.194	-0.228	-0.200
	(0.185)	(0.150)	(0.194)	(0.183)
Random Treat × Non-Participant	-0.093	-0.117	-0.039	-0.023
	(0.134)	(0.101)	(0.157)	(0.146)
Baseline Outcome		0.335***		0.126**
		(0.042)		(0.056)
Constant	-0.028	0.034	0.076	0.053
	(0.097)	(0.073)	(0.067)	(0.071)
R-squared	0.004	0.124	0.006	0.024

Notes: Regressions restricted to Random Treatment and Control. N = 920 students in 20 schools in all specifications. Robust standard errors in parentheses, clustered by school. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Omitted category is all girls in Control.

Next, I investigate treatment effect heterogeneity. Heterogeneity may occur along many dimensions, such as those defined by the variables in Table 1.1. In order to reduce dimensionality, I use the Control schools to predict endline outcomes conditional on baseline outcomes and individual-level covariates. This takes the form of regression results presented in Table 1.7. In a sense, this uses the Control group as a counterfactual to predict what would have occurred in treatment schools in the absence of treatment, conditional on variables observed at baseline. From these results, we see that being in a Scheduled Caste or Scheduled Tribe predicts approximately 0.5 standard deviations lower Educational Aspirations, as shown in Column (1). Using the predicted outcomes from this regression, I then group students into predicted outcome terciles. Low, Middle, and High predicted terciles are denoted by $\hat{\text{L}}$, $\hat{\text{M}}$, and $\hat{\text{H}}$, respectively.[9]

[9] While the table presents the coefficient estimates used to predict $\hat{\text{L}}$, $\hat{\text{M}}$, and $\hat{\text{H}}$ in both Treatment arms, I use a leave-one-out procedure suggested in Abadie, Chingos and West (2014) to predict outcome terciles for students in Control. Abadie, Chingos and West (2014) show through simulation that such a procedure

Table 1.7: Defining Predicted Outcome Terciles

	Educational Aspirations (1)	Gender Roles (2)
Elected	0.039	-0.009
	(0.113)	(0.058)
Grade 7	0.181	0.087
	(0.203)	(0.196)
Grade 8	0.106	0.166
	(0.189)	(0.164)
Scheduled Caste	-0.509**	-0.165
	(0.168)	(0.161)
Scheduled Tribe	-0.517*	-0.804
	(0.229)	(0.484)
Other Backwards Caste	-0.359	-0.272***
	(0.208)	(0.059)
Baseline Outcome	0.302***	0.027
	(0.059)	(0.057)
Constant	0.288*	0.222
	(0.145)	(0.175)
Observations	393	395
R-squared	0.164	0.052

Robust standard errors in parentheses, clustered by school. *** p<0.01, ** p<0.05, * p<0.1. Estimation restricted to students in Control schools.

Using these predicted outcome terciles, I estimate heterogeneous treatment effects with regressions of the form in Equation (1.2).

$$y_{is} = \sum_{k=1}^{3} I_k(\beta_{0k} + \beta_{1k} RandomTreat_s \times Participant_{is} + \beta_{2k} RandomTreat_s \times NonParticipant_{is}) + \epsilon_{is} \tag{1.2}$$

In this specification, I_k is an indicator for being in each predicted tercile.[10] Results for this specification appear in Table 1.8. I also estimate versions of Equation 1.2 that include baseline outcomes interacted with $\hat{\mathrm{L}}$, $\hat{\mathrm{M}}$, and $\hat{\mathrm{H}}$. Note that there are strongly negative effects for Gender Roles among Participants in Random Treatment, but only for those in the middle predicted tercile. This presents suggestive evidence of heterogeneous treatment effects for Gender Roles in Random Treatment schools, with heterogeneity defined by predicted outcome tercile. Further, while noting lack of power to detect small differences, there are no significant effects on non-participants for any predicted outcome tercile.

1.2.6.2 Effects on Networks

While I find suggestive evidence for negative treatment effects on outcomes, there is much stronger evidence for treatment effects on networks. Since we have random within-school variation in Random Treatment schools, I present reduced-form treatment effect estimates broken down by whether each node involved in the link is chosen for participation. To do this, I estimate Equation (1.3).

$$L_{ijs} = \gamma_0 + \gamma_1 Participant_{is} + \gamma_2 Participant_{js} + \gamma_3 Participant_{is} \times Participant_{js} + u_{ijs} \tag{1.3}$$

solves the overfitting bias that results from endogenous stratification.

[10] That is, $I_1 = \hat{L}$, $I_2 = \hat{M}$, $I_3 = \hat{H}$.

Table 1.8: Treatment Effect Heterogeneity by Predicted Outcome Tercile

	Education		Gender Roles	
	(1)	(2)	(3)	(4)
$\hat{\text{L}}$	-0.417**	0.061	-0.021	-0.018
	(0.191)	(0.178)	(0.152)	(0.149)
$\hat{\text{M}}$	-0.064	-0.057	0.124*	0.107
	(0.098)	(0.093)	(0.069)	(0.074)
$\hat{\text{H}}$	0.346***	0.251***	0.108	0.038
	(0.054)	(0.056)	(0.099)	(0.116)
Participant in Random Treat × $\hat{\text{L}}$	-0.323	-0.422	-0.048	-0.029
	(0.277)	(0.250)	(0.215)	(0.217)
Participant in Random Treat × $\hat{\text{M}}$	-0.090	-0.112	-0.610***	-0.582***
	(0.127)	(0.124)	(0.168)	(0.158)
Participant in Random Treat × $\hat{\text{H}}$	0.139	0.120	0.135	0.114
	(0.201)	(0.205)	(0.327)	(0.291)
Non-Participant in Random Treat × $\hat{\text{L}}$	-0.061	-0.110	0.038	0.049
	(0.198)	(0.177)	(0.183)	(0.183)
Non-Participant in Random Treat × $\hat{\text{M}}$	-0.035	-0.074	0.004	0.006
	(0.126)	(0.119)	(0.214)	(0.202)
Non-Participant in Random Treat × $\hat{\text{H}}$	-0.090	-0.098	-0.176	-0.188
	(0.146)	(0.144)	(0.238)	(0.209)
Baseline Outcome Interactions	NO	YES	NO	YES
R-squared	0.120	0.147	0.024	0.042
Test1 P-value	0.354	0.209	0.009	0.011
Test2 P-value	0.922	0.978	0.587	0.492

Notes: Regressions restricted to Random Treatment and Control. N = 920 in 20 schools in all specifications. Omitted category is all girls in Control. Robust standard errors in parentheses, clustered by school. *** p<0.01, ** p<0.05, * p<0.1. $\hat{\text{L}}$, $\hat{\text{M}}$, and $\hat{\text{H}}$ predicted from baseline variables (see Table 1.7). Baseline Outcome Interactions include interactions of Baseline Outcome with $\hat{\text{L}}$, $\hat{\text{M}}$, and $\hat{\text{H}}$. Test1 is a test of equality of the interactions of $\hat{\text{L}}$, $\hat{\text{M}}$, and $\hat{\text{H}}$ with Participant in Random Treat. Test2 is a test of equality of the interactions of $\hat{\text{L}}$, $\hat{\text{M}}$, and $\hat{\text{H}}$ with Non-Participant in Random Treat.

These results are presented in Table 1.9, where Columns (2) and (4) additionally control for baseline link values.

Results for the binary link definition show significant effects for the interaction term γ_3. That is, if both students are chosen to participate, the probability of a link is much larger than if only one is chosen. Similar patterns hold for the continuous definition in Columns (3) and (4), where we see much larger average link value if both are chosen. For the continuous link definition, we see evidence for *substitution* of links: if only one is chosen, the average link value decreases.

These results contain powerful implications for evaluation of counterfactual assignment policies. Participation in the program has a substantial effect on the identity of others with whom individuals interact. If we hypothesize that program effects diffuse through networks, then failing to account for the effect of the program *on the structure of the network itself* may lead to erroneous predictions. Accordingly, estimates that interpret reduced-form effects as the direct effects of the program will miss this important channel of change. Further, any attempt to predict outcomes under counterfactual assignments needs to account for the effect of the program on networks.

1.2.7 Reduced-Form Stylized Facts

The descriptive analysis as well as reduced-form treatment effects provide a number of stylized facts that a model must rationalize. At baseline, we see that outcomes vary substantially with observed characteristics, especially caste groupings. We further see evidence for the following six stylized facts about networks.

1. Networks are not independent of observed characteristics but rather exhibit substantial homophily.

2. Links are not symmetric but are complementary. That is, the existence/intensity of individual i's link to individual j is positively correlated with the existence/intensity of individual j's link to i.

Table 1.9: Reduced-Form Treatment Effects on Networks

Network Definition	Binary		Continuous	
	(1)	(2)	(3)	(4)
Participant (Self)	-0.031	-0.032	-0.059	-0.102**
	(0.027)	(0.021)	(0.039)	(0.040)
Participant (Alter)	0.012	0.004	0.078	0.021
	(0.015)	(0.013)	(0.044)	(0.034)
Participant (Both)	0.117**	0.096**	0.308**	0.220**
	(0.037)	(0.031)	(0.110)	(0.085)
Baseline Measure (Self)		0.251***		0.240***
		(0.014)		(0.014)
Baseline Measure (Alter)		0.125***		0.096***
		(0.010)		(0.019)
Constant	0.536***	0.309***	1.231***	0.920***
	(0.015)	(0.013)	(0.032)	(0.040)
R-squared	0.003	0.086	0.007	0.092
P-value of Test	0.041	0.082	0.048	0.163

Notes: Sample restricted to girls in Random Treatment schools. N = 19,430 in 20 schools in all specifications. Robust standard errors in parentheses, clustered by school. *** p<0.01, ** p<0.05, * p<0.1. Dependent variable is existence/intensity of link between i (self) and j (alter), as indicated by i at endline. Test is a test of significance of sum of coefficients for Participant (self), Participant (Alter), and Participant (Both), against a null that the sum is zero. Missing network data imputed via algorithm described in Appendix D.

3. Average number of links (for binary measure) and sum of link values (for continuous measure) is growing in school size.

4. Average link value is decreasing in school size, suggesting that links are not independent. Rather, this suggests that there are tradeoffs to an individual in making link decisions.

5. The program negatively affects outcomes, but these effects are insignificant.

6. The program substantially affects network links, even conditional on baseline networks.

The structural model that follows provides features that account for all of these patterns.

1.3 Peer Effects Model

Here, I describe the model of peer effects. This provides the key identification result for the peer effects model, which posits network endogeneity as an omitted-variable problem (see, e.g., Goldsmith-Pinkham and Imbens, 2013; Hsieh and Lee, 2016), while also allowing for non-linear peer effects (Carrell, Sacerdote and West, 2013). Conditional on observing the confounding omitted variables, the parameters of the peer effects model are identified even in the presence of network endogeneity.

1.3.1 The Problem of Endogenous Networks

The peer effects model begins with a reduced-form version of the standard linear-in-means model (see, e.g., Manski, 1993):

$$y_{is} = \alpha_0 + \alpha_1 P_{is} + \alpha_2 \bar{P}_{is} + u_{is} \tag{1.4}$$

where y_{is} is some outcome, such as test scores or, in the data here, Educational Aspirations.[11] As discussed in Carrell, Sacerdote and West (2013), with some assumptions the canonical model can be rewritten as Equation (1.4).[12] P_{is} is an indicator for individual i in school s being chosen to participate in the program, and $\bar{P}_{is}$ is individual i's peer group mean participation. The variable u_{is} is unobserved. Therefore, from Equation (1.4), we see that expected outcome y_{is} is a linear function of a student's participation status and the participation status of her peers, along with an additional additive unobserved component.[13]

Equation (1.4) requires a definition of the peer group mean variable $\bar{P}_{is}$. This in turn requires the choice of how to weight peers. Suppose that in each school s we observe a matrix $\mathbf{G_s}$ of directed links between individuals i and j, where element (i,j) corresponds to individual i's link to j. Now,

$$\bar{P}_{is} = \sum_{j \neq i} \frac{w_{ijs}(\mathbf{G_s})}{\sum_{k \neq i} w_{iks}(\mathbf{G_s})} P_{js} \tag{1.5}$$

where $w_{ijs}(\mathbf{G_s})$ is a function from the link matrix to define the weight for the link between individuals i and j. For example, when links are binary, then $w_{ijs}(\mathbf{G_s}) \in \{0,1\}$ and thus $\bar{P}_{is}$ is merely the fraction of an individual's peers who are also chosen to participate. If link values are continuous, then $\bar{P}_{is}$ may weight "closer" or "stronger" links more. For purposes here, with directional links g_{ijs} and g_{jis} defining the links between individuals i and j, I take as given the weighting function $w_{ijs}(\mathbf{G_s}) = g_{ijs} + g_{jis}$. See Appendix C for a fuller discussion of the issue of weighting.

I augment this model by decomposing the error term u_{is} in a manner similar to Goldsmith-Pinkham and Imbens (2013) and Hsieh and Lee (2016), who effectively include the unob-

[11] Blume et al. (2015) provide micro-foundations for this model as well as a generalization of various identification results derived since Manski (1993).

[12] Note that the bulk of the literature focuses on estimating an endogenous peer effect, which is not the focus of the model here.

[13] This model effectively assumes that peer influence is characterized by the peer group mean. Some recent studies, in contrst, have presented evidence of the importance of the variance in peer ability (Booij, Leuven and Oosterbeek, 2016; Lyle, 2009).

served variable a_{is}. In contrast to those papers, however, I follow the suggestion by Bramoullé (2013) to include a peer effect in the unobserved variable, which amounts to including $\bar{a}_{is}$ as an additional regressor in the structural peer effects model.[14]

Let $u_{is} = \alpha_3 a_{is} + \alpha_4 \bar{a}_{is} + v_{is}$. Accordingly, Equation (1.4) becomes Equation (1.6).

$$y_{is1} = \alpha_0 + \alpha_1 P_{is} + \alpha_2 \bar{P}_{is} + \alpha_3 a_{is} + \alpha_4 \bar{a}_{is} + v_{is} \tag{1.6}$$

With this formulation, network endogeneity biases peer effects estimates whenever $\bar{P}_{is}$ is correlated with either a_{is} or $\bar{a}_{is}$. That is, if one estimates Equation (1.4) without controlling for a_{is}, estimates of α_2 will be biased due to correlation between $\bar{P}_{is}$ and u_{is} (which includes $(a_{is}, \bar{a}_{is})$).

As an example for when $\text{cov}(\bar{P}_{is}, a_{is}) \neq 0$, suppose that a_{is} is unobserved academic ability, and this unobserved ability is positively associated with outcome y_{is}. Endogeneity arises when a_{is} *also* plays a part in the network formation process, such as if those with higher ability also are more likely to link with participants. Therefore, those with higher a_{is} will tend to have more of their links be with participants, bringing about positive correlation between $\bar{P}_{is}$ and a_{is}.[15] Note that this endogeneity may arise even when P_{is} is exogenous, such as the case when participation is assigned randomly. That is, even with random assignment, estimation that does not account for unobserved a_{is} may be biased in the presence of endogenous network formation.

In addition to accounting for network endogeneity, I further allow for the possibility of non-linear peer effects. As in Carrell, Sacerdote and West (2013), these non-linear peer effects account for the fact that peer means $\bar{P}_{is1}$ and $\bar{a}_{is}$ may affect different types of individuals differently. This is accounted for by the variables $I_{isk}(z_{is}), k = 1, ..., K$, which define a set of K indicators for being in different categories of the population. The partition could

[14] $\bar{a}_{is}$ is defined analogously to $\bar{P}_{is}$: $\bar{a}_{is} = \sum_{j \neq i} \frac{w_{ijs}(\mathbf{G_s})}{\sum_{k \neq i} w_{iks}(\mathbf{G_s})} a_{js}$. Other peer-group mean variables are defined similarly.

[15] It is not sufficient that those with higher ability link more (or less) with all students. Endogeneity arises because ability leads to differential valuation of network links based on ability.

be defined by grade level, gender, baseline outcome, or any other function of exogenous characteristics z_{is}. With this additional step, Equation (1.6) becomes Equation (1.7).

$$y_{is} = \sum_{k=1}^{K} I_{isk} \left(\alpha_{0k} + \alpha_{1k} P_{is} + \alpha_{2k} \bar{P}_{is} + \alpha_{3k} a_{is} + \alpha_{4k} \bar{a}_{is} \right) + v_{is} \tag{1.7}$$

Non-linear peer effects are captured by the coefficients a_{2k} and a_{4k} varying with different values of k.[16]

Finally, in order to increase power, I also estimate a version of Equation (1.7) that includes baseline outcomes. This specification is defined by Equation (1.8), where outcomes are now indexed by time t, with $t = 1$ corresponding to endline and $t = 0$ to baseline.

$$y_{is1} = \sum_{k=1}^{K} I_{isk} \left(\alpha_{0k} + \alpha_{1k} P_{is} + \alpha_{2k} \bar{P}_{is1} + \alpha_{3k} a_{is} + \alpha_{4k} \bar{a}_{is1} + \alpha_{5k} y_{is0} + \alpha_{6k} \bar{y}_{is0} \right) + v_{is1} \tag{1.8}$$

1.3.2 Identification Results for the Peer Effects Model

With the outcome equation formulated as in Equation (1.7), identification is straightforward. Define the parameter vector $\alpha = (\alpha_{01}, ..., \alpha_{41}, ..., \alpha_{0K}, ..., \alpha_{4K})$. Let N_s be the number of students in school s. For each s, define $\mathbf{P_s} = (P_{1s}, ..., P_{N_s s})'$ and $\mathbf{A_s} = (a_{1s}, ..., a_{N_s s})'$. Conditional on independence of observations across schools as well as exogeneity of participation ($\mathbf{P_s}$), the unobserved confounders ($\mathbf{A_s}$), and the network ($\mathbf{G_s}$), α is identified. This result is formalized in Proposition 1.1.

Proposition 1.1. *Suppose that*

1. $(P'_{is}, \bar{P}_{is}, a_{is}, \bar{a}_{is}) \perp\!\!\!\perp (P_{jt}, \bar{P}_{jt}, a_{jt}, \bar{a}_{jt}) \ \forall \ s \neq t$.

2. $\alpha \in \mathbf{\Theta}_\alpha \subset \mathbb{R}^{5K}$, *where* $\mathbf{\Theta}_\alpha$ *is compact.*

3. $(P_{is}, \bar{P}_{is}, a_{is}, \bar{a}_{is}) \in \mathbf{X} \subset \mathbb{R}^4$, *where* $\mathbf{X}$ *is compact.*

[16] This model allows for non-linear *direct* effects (captured by α_{1k} and α_{3k}) as well as non-linear *peer* effects (captured by α_{2k} and α_{4k}).

4. $\mathbb{E}[v_{is}|\mathbf{P_s}, \mathbf{G_s}, \mathbf{A_s}] = 0 \; \forall \; j$

5. $\mathbb{E}[D_{is}D_{is}']$ *is of rank* $5K$*, where* $D_{is} = [I_{is1}(1, P_{is}', \bar{P}_{is}, a_{is}, \bar{a}_{is})', ..., I_{isK}(1, P_{is}', \bar{P}_{is}, a_{is}, \bar{a}_{is})'] \in \mathbb{R}^{5K}$.

Then the parameter vector α *of Equation (1.7) is identified as* $s \to \infty$*.*

Proof. This is a standard OLS result other than the fact that it relies upon consistent estimates of a_{is}. □

This model generalizes both the standard linear-in-means model as well as well as the more general model used by Carrell, Sacerdote and West (2013). The authors of that paper essentially assumed that $\alpha_{3k} = \alpha_{4k} = 0$ for all k. The standard linear-in-means model typically further assumes that $K = 1$ and thus $I_{is1} = 1 \; \forall \; i, s$, implying no non-linear effects. Accordingly, the identification result in Equation (1.1) states weaker conditions than those previously used in the literature on peer effects.

Finally, I note that identification and thus consistent estimation in the presence of network endogeneity depends crucially on an initial estimate of unobserved a_{is}. This estimate is obtained from estimation of the network-formation process, which is described in the next section. Conditional on a_{is} and given the assumptions of Proposition 1.1, we can recover the true parameters of the peer effects model.[17]

1.3.3 Relation to Other Models

The peer effects model here combines two approaches that have received substantial attention in the statistics and econometrics literature. First, I specify arbitrary latent characteristics a that must be accounted for. Second, conditional on these latent characteristics, the model posits a parametric control function approach. These twin approaches allow for

[17] Average school size in the data is 44 girls, and thus we have 86 data points with which to estimate a_{is} for each i, corresponding to i's 43 decisions to link to others and the 43 others' decisions to link to i. If, despite this, the prior estimation returns noisy but unbiased estimates of a_{is}, this should induce attenuation in estimates of α.

identification of the parameters of the peer effects model in the presence of certain types of network endogeneity.

First, I posit a as an unobserved, "latent" characteristic, making this model similar to the "latent space" models described in Jackson (2014). Such models posit that unobserved "latent" characteristics play a part in the process being modeled. As Jackson (2014) discusses, a key feature of such models is that the latent characteristic may be *any* unobserved—and possibly difficult-to-measure—characteristic, such as "ability" or "ambition."[18] Goldsmith-Pinkham and Imbens (2013) models the latent characteristic as a single binary variable, while Hsieh and Lee (2016) allow for continuous multi-dimensional unobservables.

Further, conditional on these latent characteristics a, identification of the model's parameters is achieved via a parametric control function approach. Rearrangement of Equation (1.7) shows this.

$$y_{is1} = \sum_{k=1}^{K} I_{isk}\left(\alpha_{0k} + \alpha_{1k}P_{is} + \alpha_{2k}\bar{P}_{is1}\right) + \sum_{k=1}^{K} I_{isk}\left(\alpha_{3k}a_{is} + \alpha_{4k}\bar{a}_{is1}\right) + v_{is1}$$

$$y_{is1} = \sum_{k=1}^{K} I_{isk}\left(\alpha_{0k} + \alpha_{1k}P_{is} + \alpha_{2k}\bar{P}_{is1}\right) + f(a_{is}, \bar{a}_{is}, z_{is}) + v_{is1} \tag{1.9}$$

$$y_{is1} = \sum_{k=1}^{K} I_{isk}\left(\alpha_{0k} + \alpha_{1k}P_{is} + \alpha_{2k}\bar{P}_{is1}\right) + u_{is1} \tag{1.10}$$

The control function is $f(a_{is}, \bar{a}_{is}, Z_{is})$ in Equation (1.9). Endogeneity arises because P_{is} and $\bar{P}_{is}$ may depend on a_{is} and $\bar{a}_{is}$. This implies correlation between these regressors and u_{is1}, leading to biased estimates of α_{1k} and α_{2k} if estimating Equation (1.10). On the other hand, estimating the control function f in Equation (1.9) allows for identification in the presence of this endogeneity. That is, the parameters of the model are identified under strictly weaker exogeneity assumptions than are typically assumed in the literature. For example, Carrell, Sacerdote and West (2013) effectively assume exogeneity of u_{is1} in Equation (1.10), while the method here only requires the exogeneity of v_{is} in Equation (1.9).

[18] "Latent space" models have also been heavily used in industrial organization, for example in Berry, Levinsohn and Pakes (1995), in their pathbreaking methodology for demand estimation.

Identification in the presence of endogeneity via control functions has found wide application in applied econometrics, and the model here follows in this tradition. As pointed out by Bramoullé (2013), the model in Goldsmith-Pinkham and Imbens (2013) is similar in spirit to the canonical Heckman selection model (Heckman, 1979), which itself uses a parametric control function approach to identification. Employment of control functions to account for unobserved heterogeneity has also found widespread application in industrial organization, particularly in the estimation of production functions (Ackerberg, Caves and Frazer, 2015; Levinsohn and Petrin, 2003; Olley and Pakes, 1996).

1.4 A Structural Model of Network Formation

The prior section showed that, conditional on the observed network and unobserved variables a_{is}, the parameters of the peer-effects model are identified. This section demonstrates that these unobeserved variables a_{is} are identified through observation of the network-formation process. Therefore, after they are recovered through the network-formation estimation procedure, these variables can be plugged in to provide for consistent estimation of the peer-effects parameters.

1.4.1 Simple Model

To fix ideas and intuition, I develop a simple version of the network-formation model. This simple model sets aside the unobserved variables a_{is} that will be added into the model later. Through this, we develop intuition behind equilibrium results, the instrumentation strategy, and the conditions for identification.

1.4.1.1 Players, Strategy Space, and Utility

For a given school s, there are N_s players in the network formation game. N_s is assumed to be determined exogenously. In the context here, N_s is the number of girls in a given school s in grades 6-8.

Each player i in school s chooses whether to be linked to each of the other $N_s - 1$ players. More formally, each player i in school s chooses a vector of actions $g_{is} \in \mathbb{R}_{+}^{N_s-1}$. Importantly, in a deviation from the bulk of the theoretical network-formation literature, link intensity is continuous: $g_{ijs} \in [0, \infty)$.[19] Further, individuals make these choices subject to a total effort constraint as spelled out in Assumption 1.1. Each individual's objective is to maximize utility subject to this constraint.

Assumption 1.1. *For each $i = 1, ...N_s$, $\sum_{j \neq i} c_{ijs} g_{ijs} \leq M_{is}$, where c_{ijs} is the cost to individual i of forming a link with j and M_{is} is individual i's endowment. Further, $M_{is} \in [\underline{M}, \overline{M}] \subset \mathbb{R}_{++}$ and $c_{ijs} \in [\underline{c}, \overline{c}] \subset \mathbb{R}_{++}$.*

The budget constraint serves two purposes in the model.[20] First, it imposes a structured way in which individuals trade off the costs and benefits of different linking strategies. If individual i's constraint is binding and she chooses to increase g_{ijs} (her link to j), then she must decrease some g_{iks} (her link to another student k). Second, M_{is} may vary across students and may depend on observed or unobserved characteristics. Accordingly, M_{is} allows for out-degree heterogeneity: individuals with higher M_{is} have a higher effort endowment and thus will tend to have more out-links in equilibrium, conditional on other variables in the model. Finally, note that the lower bound on cost implicitly imposes the restriction that network size is bounded above for each individual: even as the size of a person's school grows infinitely, the sum of links can only grow so much: $\sum_{j \neq i} g_{ijs} \leq \frac{M_{is}}{\underline{c}}$. Compact support of M_{is} implies further that network size is bounded above across individuals.

Utility for individual i in school s is a function of the realized network $\mathbf{G_s}$ as well as exogenous characteristics of all students in school s, $\mathbf{X_s}$, where $\mathbf{X_s} = (X'_{1s}, ...X'_{N_s s})'$. Following prior models (Badev, 2013; Mele, 2010, e.g.,), I assume that the utility of links is additive. Similar to these models, I assume that individuals derive different utilities depending upon

[19] Exceptions are found in Baumann (2016), Bloch and Dutta (2009), and Rogers (2006). Jackson (2008) briefly mentions models of this type in a section titled "Weighted Network Formation."

[20] Budget constraints are rare in models of network formation, but have found application in both continuous (Baumann, 2016; Bloch and Dutta, 2009) and discrete (Boucher, 2015) models of network formation.

how "mutual" their links are. The utility to individual i of a network $\mathbf{G_s}$ is given in Equation (1.11).

$$\begin{aligned} U_{is}(\mathbf{G_s}, \mathbf{X_s}) &= \sum_{j \neq i} u_{ijs} \\ &= \sum_{j \neq i} g_{ijs}^{\alpha} g_{jis}^{\beta} e^{f(X_{is}, X_{js})} \end{aligned} \tag{1.11}$$

The utility to individual i from his link to j depends upon both his linking strategy g_{ijs} and on j's linking strategy via g_{jis}. The Cobb-Douglas function imposes complementarity in linking strategies. Further, the functional form implies that all links are marginally valuable, except when $g_{jis} = 0$. Hence, in the absence of a budget constraint, all individuals would choose to be maximally linked to all others.

Assumption 1.2. *The following restrictions hold:*

1. *X_{is} and $f()$ are bounded in $\mathbb{R}^k$ and $\mathbb{R}$, respectively.*

2. $0 < \beta < (1 - \alpha) < 1$

Assumption 1.2 imposes additional structure on the utility function, and these assumptions have important implications for equilibrium. The bulk of the literature on network formation with continuous link values assumes that the utility function is *convex* in own strategy, which would be implied here if $\alpha > 1$ (see, e.g., Bloch and Dutta, 2009). As pointed out by Boucher (2015), this leads to equlibria in which actors form few strong links, and the equilibrium set is qualitatively similar to the case when strategy sets are discrete. Further, the assumption that $\beta < (1 - \alpha)$ contrasts with Baumann (2016), who assumes $\beta = (1 - \alpha)$, and this different assumption leads to different sets of equilibrium strategy profiles.

Before discussing equilibrium, I note that the model is limited in two important ways. First, utility from given links depends *only* on the link between those two individuals as well as their characteristics. Importantly, the utility to i of linking to j does *not* depend upon j's links, other than his link to i. Thus, this model does not allow for utility from linking to popular individuals. Conversely, it does not allow for congestion externalities, whereby a link to a given individual is less valuable when that individual has more links. Through the budget constraint, however, the model *does* allow for tradeoffs between links.

Second, and in contrast to the structure discussed in Bramoullé (2013) and Blume et al. (2015), individuals do not consider final outcomes y_{is} in making their linking decisions. This assumption is made more plausible in educational contexts by the findings in Carrell, Sacerdote and West (2013), who show that Air Force members tend to choose peers by homophily. In such a context, at least for parts of the population, a homophilic linking strategy would tend to lead to lower academic outcomes.

1.4.1.2 Equilibrium

Here, I provide results showing equilibrium existence and uniqueness. In contrast to models with discrete action spaces, the continuity of the model allows for the use of Nash equilibrium rather than pairwise stability. Further, while acknowledging the existence of many corner equilibria, I show that there exists a single equilibrium in which all individuals are linked.

As spelled out above, each individual chooses a vector of links g_{is} to maximize his utility subject to others' linking decisions. Proposition 1.2 provides the paper's primary existence result. Existence is guaranteed by the concavity of the game, a result that dates back at least to Rosen (1965). However, the Nash Equilibrium is not unique: it is possible to have an equilibrium with any combination of $g_{ijs} = g_{jis} = 0$. For example, there exists an equilibrium in which each person is connected only to one other person, on whom he exhausts his entire endowment of effort. Further, a completely empty network, in which $g_{ijs} = g_{jis} = 0$ for all $i, j \neq i$ is an equilibrium.

Accordingly, to refine the set of equilibria, I define a *strictly positive equilibrium* as a Nash equilibrium in which each person's strategy profile exhibits strictly positive links. That is, for a strategy profile to be a strictly positive equilibrium, it must be a Nash equilibrium and $g_{ijs} > 0$ for every $i, j \neq i$.

Proposition 1.2. *There exists a Nash equilibrium for the network-formation game. Further, there exists a strictly positive equilibrium.*

Proof. See Appendix A. □

A necessary condition for a strictly positive equilibrium is that the following first-order conditions hold:

$$\frac{\partial U_{is}}{\partial g_{ijs}} = \alpha g_{ijs}^{\alpha-1} g_{jis}^{\beta} e^{f(X_{is}, X_{js})} - c_{ijs}\lambda_{is} = 0 \;\; \forall\, i,\; j \neq i \tag{1.12}$$

$$\frac{\partial U_{is}}{\partial \lambda_{is}} = M_{is} - \sum_{j \neq i} c_{ijs} g_{ijs} = 0 \;\; \forall\, i \tag{1.13}$$

Importantly, there is only one interior equilibrium, as stated in Proposition 1.3. Intuitively, uniqueness derives from the concavity of the network-formation game. The result states that there is a *unique* solution to the First-Order Conditions in Equations (1.12) and (1.13) that characterize the strictly positive equilibrium.

Proposition 1.3. *The strictly positive equilibrium of the game is unique.*

Proof. See Appendix A. □

This uniqueness result is quite important for estimation and simulation. First, identification and estimation proceed by assuming we observe the network in this unique state. Second, conditional on the parameters of the model, we can simulate counterfactuals by finding any solution to these conditions, with the knowledge that no others exist.

1.4.1.3 Relation to Potential Games

As further justification for focusing attention on the strictly positive equilibrium, here I relate the game as developed to the theory of potential games. Several prior papers dealing with identification and estimation of network formation games have leveraged results from the potential games literature (Badev, 2013; Mele, 2010). I show that a special case of the game developed here is a potential game. I further demonstrate that, in this special case, the potential function allows for refinement of the set of equilibrium strategy profiles to include only the strictly positive equilibrium.

In general, the network formation game is not a potential game. To see this fact, note that

$$\frac{\partial^2 U_{is}}{\partial g_{ijs} \partial g_{jis}} = \alpha\beta g_{ijs}^{\alpha-1} g_{jis}^{\beta-1} e^{f(X_{is}, X_{js})} \tag{1.14}$$

$$\frac{\partial^2 U_{js}}{\partial g_{jis} \partial g_{ijs}} = \alpha\beta g_{jis}^{\alpha-1} g_{ijs}^{\beta-1} e^{f(X_{js}, X_{is})} \tag{1.15}$$

In a simpler setting, Theorem 4.5 in Monderer and Shapley (1996) states that a sufficient and necessary condition for a continuous game to be a potential game is that the partial derivatives in Equations (1.14) and (1.15) are equal, which in general does not hold here.[21] However, there exists a special case in which this symmetry condition does hold. In this special case, the network formation game is a potential game with the potential function defined in Proposition 1.4.

Proposition 1.4. *If $\beta = \alpha$ and $f(X_{is}, X_{js}) = f(X_{js}, X_{is})$, then the network formation game is an exact potential game with potential function defined as*

$$P(\mathbf{G_s}, \mathbf{X_s}) = \frac{1}{2} \sum_{i=1}^{N_s} \sum_{j \neq i} (g_{ijs} g_{jis})^{\alpha} e^{f(X_{is}, X_{js})}$$

Proof. See Appendix A. □

In this special case of the game, equilibrium results follow directly from known results of

[21] I note that they proved this fact when strategy sets are intervals of the real line. In my setting, strategy sets are compact subsets of $\mathbb{R}^{N_s - 1}$. Accordingly, I do not rely on their theorem to prove Proposition 1.4.

potential games. Importantly, Monderer and Shapley (1996) note that the potential function can be used for equilibrium refinement, in a way that is useful for the task at hand. They show that the set of potential maximizers is a subset of the set of Nash equilibria.

Proposition 1.5. *If $\beta = \alpha$ and $f(X_{is}, X_{js}) = f(X_{js}, X_{is})$, then the unique strictly positive equilibrium is the unique maximizer of the potential function $P(\mathbf{G_s}, \mathbf{X_s})$.*

Proof. See Appendix A. □

Corollary 1.1. *The unique strictly positive equilibrium is efficient in that it maximizes total utility.*

Proof. From the definition of the potential function, $P(\mathbf{G_s}, \mathbf{X_s}) = \frac{1}{2}\sum_{i=1}^{N_s} U_{is}(\mathbf{G_s}, \mathbf{X_s})$. Therefore, the set of values that maximizes the potential function also maximizes the sum of utilities, and the result follows from Proposition 1.5. □

Proposition 1.5 provides an important result for the game at hand. While the set of Nash equilibria is quite large, the set of strategy profiles that maximize the potential function is a singleton, containing only the strictly positive equilibrium strategy profile. Further, Corollary 1.1 demonstrates that, in this special case, the strictly positive equilibrium maximizes the sum of individuals' utility. Note that this is a quite strong efficiency result, whereby the strictly positive equilibrium maximizes total utility among all sets of feasible strategies, which is a much larger set than the set of Nash equilibria. These results for potential games provide additional support for the focus on the strictly positive equilibrium in the empirical analysis that follows.

1.4.2 Identification Results for the Simple Model

The prior subsection showed that there exists a unique strictly positive Nash equilibrium. Identification proceeds by assuming that we observe S networks in this equilibrium state. In this subsection, I provide conditions under which parameters of the network formation game are identified. In particular, Assumption 3 states what is observed. I note that this assumption allows for observations to be arbitrarily dependent *within* schools.

Assumption 1.3. *For each $s = 1, ..., S$, $i = 1, ..., N_s$, we observe a vector of characteristics and links $(X'_{is}, g'_{is}) \in \mathbf{X} \times \mathbf{G}$, where $\mathbf{X} \subset \mathbb{R}^m$ and $\mathbf{G} \subset \mathbb{R}^{N_s-1}$ are compact, $m = dim(X_{is})$, and N_s is the number of agents in school s. Further, $(X'_{is}, g'_{is}) \perp\!\!\!\perp (X'_{jt}, g'_{jt})\ \forall\ s \neq t$.*

1.4.2.1 Identification Arguments for Networks

Before proceeding to identification results, I take a slight detour to discuss identification in the context of networks. Dependence among observed network links complicates asymptotics. In the model developed here, dependence has two sources. First, since utility depends upon the mutual-ness of links, individual i's link choice to j depends on j's choice to i. So, g_{ijs} depends on g_{jis}, where $j \neq i$. Second, the budget constraint imposes dependence among all of an individual's links. That is, g_{ijs} depends on g_{iks}, where $j, k \neq i$. Accordingly, we require identification arguments that account for these cross-sectional dependencies.

To account for these dependencies, identification results in network-formation models have taken two different strategies, both of which I employ here. Leung (2015) refers to these two strategies as "many market" and "large market" asymptotics. First, many market asymptotics depend upon observation of a number of different networks. That is, in our context, identification is achieved as S, the number of schools, approaches infinity. Such arguments can be employed to identify parameters that are common across networks.

In contrast, identification of parameters that are only observed within a single market requires observation of arbitrarily large networks. As discussed in Graham (2014*a*), for a network with N_s agents, the econometrician observes $N_s - 1$ linking decisions per agent. Importantly in our context, we need to identify individual-specific parameters a_{is} that are only observed within a single school. Identification of these parameters leverages such large market asymptotics, where parameters are identified as the size of the network s—which contains individual i—grows. Such asymptotics are non-standard, since the dimension of the parameter vector is also increasing.

1.4.2.2 Instrumentation Strategy and Identification

The network formation model as spelled out above has two sources of endogeneity, for which I employ two distinct strategies. First, I difference out endogenous variables that depend only on i. Second, to control for endogeneity of individual j's network choice, I employ a budget set instrument, whereby exogenous variation is obtained via variation in the utility of potential links. I then show that, conditional on appropriate exogeneity assumptions and rank conditions, crucial parameters of the network-formation model are identified.

Before proceeding to results, I rearrange the first-order conditions and redefine some variables. First, Equation (1.12) becomes Equation (1.16) and then Equation (1.17).[22]

$$\log g_{ijs} = \frac{\log \alpha}{1-\alpha} + \frac{\beta}{1-\alpha} \log g_{jis} + \frac{f(X_{is}, X_{js})}{1-\alpha} - \frac{\log \lambda_{is}}{1-\alpha} - \frac{\log c_{ijs}}{1-\alpha} \tag{1.16}$$

$$\tilde{g}_{ijs} = \tilde{\alpha} + \tilde{\beta}\tilde{g}_{jis} + \tilde{f}(X_{is}, X_{js}) - \tilde{\lambda}_{is} - \tilde{c}_{ijs} \tag{1.17}$$

Importantly, the parameters α and β are subsumed into a composite parameter $\frac{\beta}{1-\alpha}$, identified hereafter as $\tilde{\beta}$. Additionally, assume the following functional form:

$$\tilde{f}(X_{is}, X_{js}) = \gamma_1 X_{is} + \delta_1 X_{is} X_{js} + \gamma_3 X_{js} \tag{1.18}$$

In the data as described above, all X_{is} are binary variables. Accordingly, homophily corresponds to the coefficient δ_1 being positive (and possibly γ_1 and γ_3 being negative). Substitution and rearrangement of terms yields the following:

$$\tilde{g}_{ijs} = \tilde{\beta}\tilde{g}_{jis} + (\tilde{\alpha} + \gamma_1 X_{is} - \tilde{\lambda}_{is}) + \delta_1 X_{is} X_{js} + \gamma_3 X_{js} - \tilde{c}_{ijs} \tag{1.19}$$

[22] To get from Equation (1.16) to Equation (1.17), define and substitute $\tilde{g}_{ijs} = \log g_{ijs}$, $\tilde{\alpha} = \frac{\log \alpha}{1-\alpha}$, $\tilde{f}(X_{is}, X_{js}) = \frac{f(X_{is}, X_{js})}{1-\alpha}$, $\tilde{\lambda}_{is} = \frac{\log \lambda_{is}}{1-\alpha}$, and $\tilde{c}_{ijs} = \frac{\log c_{ijs}}{1-\alpha}$.

The econometric issue is to identify and estimate the parameters of Equation (1.19).

Identification is complicated due to two sources of endogeneity in Equation (1.19). First, $\tilde{\lambda}_{is}$, which identifies the (log) shadow value of additional effort endowment, necessarily depends upon $\tilde{c}_{ijs}$, the cost of linking. Second, whenever $\tilde{\beta} > 0$, $\tilde{g}_{jis}$ depends upon $\tilde{g}_{ijs}$, which depends on $\tilde{c}_{ijs}$. I solve these issues by using two different strategies.

The first strategy leverages the "panel" nature of the data by applying a standard differencing-out method. However, instead of the standard two dimensions of individuals i and time t, here we have two dimensions "out" i and "in" j. For all variables in Equation (1.19), perform a "within i" transformation. That is, define $\bar{g}^i_{ijs} = \frac{1}{N_s - 1} \sum_{k \neq i} \tilde{g}_{iks}$ and $\dot{g}^i_{ijs} = \tilde{g}_{ijs} - \bar{g}^i_{ijs}$. Other variables are defined similarly, leading to Equation (1.20).

$$\dot{g}^i_{ijs} = \tilde{\beta}\dot{g}^i_{jis} + \delta_1 X_{is}\dot{X}^i_{js} + \gamma_3 \dot{X}^i_{js} - \dot{c}^i_{ijs} \tag{1.20}$$

This transformation eliminates all terms that vary only with i, including the necessarily endogenous term $\tilde{\lambda}_{is}$.

Second, I employ novel instruments for the necessarily endogenous $\dot{g}^i_{jis}$ terms. The instrument relies upon tradeoffs between different linking strategies, which in turn relies upon the non-dyadic structure of the network formation model. Intuitively, due to the budget constraint, individual j's linking decision to i depends upon his alternative options for links. That is, it depends upon the utility he derives from linking to *other* individuals k, where $k \neq i$, which in turn depends upon k's characteristics. Crucially, the instrument works through the budget constraint and thus the shadow value of effort.

Simple algebra shows how these instruments are relevant. First, take the mirror image of Equation (1.19), replacing i with j and j with i, leading to Equation (1.21).

$$\tilde{g}_{jis} = \tilde{\beta}\tilde{g}_{ijs} + (\tilde{\alpha} + \gamma_1 X_{js} - \tilde{\lambda}_{js}) + \delta_1 X_{js} X_{is} + \gamma_3 X_{is} - \tilde{c}_{jis} \tag{1.21}$$

Next, perform the "within i" transformation, leading to Equation (1.22).

$$\dot{g}^i_{jis} = \tilde{\beta}\dot{g}^i_{ijs} + \gamma_1 \dot{X}^i_{js} - \dot{\lambda}^i_{js} + \delta_1 X_{is}\dot{X}^i_{js} - \dot{c}^i_{jis} \tag{1.22}$$

The terms on the right-hand side of Equation (1.22) suggest instrument candidates. However, $\dot{g}^i_{ijs}$ is the dependent variable in Equation (1.20) and thus necessarily depends on $\tilde{c}_{ijs}$. Further, $\dot{X}^i_{js}$ and $X_{is}\dot{X}^i_{js}$ are on the right-hand side of that same equation and thus not excludable. Accordingly, instruments must come through the term $\dot{\lambda}^i_{js}$.

Relevant instruments are revealed by decomposing the term $\dot{\lambda}^i_{js}$. This shows that

$$\dot{\lambda}^i_{js} = \tilde{\lambda}_{js} - \frac{1}{N_s - 1}\sum_{k \neq i} \tilde{\lambda}_{ks} = \tilde{\lambda}_{js} - \frac{1}{N_s - 1}\sum_{k} \tilde{\lambda}_{ks} + \frac{1}{N_s - 1}\tilde{\lambda}_{is} \tag{1.23}$$

The middle term is constant for all i and j within the school s. However,

$$\tilde{\lambda}_{js} = \frac{1}{N_s - 2}\sum_{k \neq i, k \neq j} \left(-\tilde{g}_{jk} + \tilde{g}_{jk}\tilde{\beta} + X_{ks}\gamma_1 + X_{js}X_{ks}\delta_1 - c_{kj}\right) \tag{1.24}$$

$$\tilde{\lambda}_{is} = \frac{1}{N_s - 2}\sum_{k \neq i, k \neq j} \left(-\tilde{g}_{ik} + \tilde{g}_{ik}\tilde{\beta} + X_{ks}\gamma_1 + X_{is}X_{ks}\delta_1 - c_{kj}\right) \tag{1.25}$$

Equations (1.24) and (1.25) motivate the use of the following instruments:

1. $\frac{1}{N_s-2}\sum_{k \neq i, k \neq j} X_{ks}$

2. $\frac{1}{N_s-2}\sum_{k \neq i, k \neq j} X_{js}X_{ks}$

3. $\frac{1}{N_s-2}\sum_{k \neq i, k \neq j} X_{is}X_{ks}$

These instruments are the mean characteristics of individuals other than i and j within school s, as well as those characteristics interacted with i's and j's characteristics.

To provide intuition for these instruments, I employ a brief example. Suppose there are three individuals in a given school: i, j, and k. Students come in two types: Wolverines and Spartans, and variable X is an indicator for being a Wolverine. Wolverine students exhibit strong homophily ($\delta_1 > 0$). Suppose i and j are both type Wolverines. Variation in k's type clearly affects i and j's link decisions to each other: if k is also a Wolverine, then both i and j will link more to k than if k is a Spartan. Due to the budget constraint, linking more to k necessitates that they link less to each other. Accordingly, variation in characteristics of *other* students serves as a relevant instrument in determining i's and j's linking strategies toward each other.

Now that relevance has been established, Assumption 1.4 provides the primary excludability assumption. This assumes mean independence of unobserved costs from *all* covariates, both those of the two individuals involved with the specific link and others. Independence of unobserved costs from all covariates is necessary for the instruments discussed above to be valid.

Assumption 1.4. $\mathbb{E}[\log c_{ijs}|X_{ks}] = 0 \ \forall \ k$.

Assumption 1.5. $(\tilde{\beta}, \delta_1', \gamma_3') \in \Theta$*, a compact subset of* $\mathbb{R}^{2m+1}$*, where* $m = dim(X_{is})$.

Proposition 1.6. *Define* $z_{ijs} = [X_{is}\dot{X}^i_{js}, \dot{X}^i_{js}, \frac{1}{N_s-2}\sum_{k \neq i,j}[X_{ks}, X_{is}X_{ks}, X_{js}X_{ks}]]$ *and* $b_{ijs} = [\dot{g}^i_{jis}, X_{is}\dot{X}^i_{js}, \dot{X}^i_{js}]$. *Given Assumptions 1.3, 1.4, and 1.5,* $(\tilde{\beta}, \delta_1', \gamma_3')$ *is identified if* $\mathbb{E}[z'_{ijs}b_{ijs}]$ *is of rank* $2m+1$.

Proof. See Appendix A. □

The simple model's main identification result is stated in Proposition 1.6. I note that, due to the "within i" transformation, parameters for terms that vary only with i are not identified. Importantly, λ_{is}, $\tilde{\alpha}$, and γ_1 are not identified, but this amounts to non-identification of the scale of each individual's utility. In contrast, parameters that identify the utility *tradeoffs* that i makes in her linking decisions are identified. The parameters $\tilde{\beta}$, δ_1, and γ_3 identify these relative tradeoffs.

There are at least two situations in which identification fails the hypotheses of Proposition 1.6. First, if δ_1 and γ_3 are both zero, then the constructed instruments are irrelevant, since then the X characteristics are irrelevant to the link-formation process. Second, the instruments may be collinear with the exogenous regressors. Importantly, if X is an indicator for treatment that is assigned by school in a randomized trial the instruments will be collinear; that is, for a given school s, $X_{is} = X_{js} = X_{ks} \; \forall \; i, j, k$. In both situations, the rank condition stated in Proposition 1.6 fails. Identification, therefore, requires that exogenous characteristics vary *within* schools and that these exogenous characteristics are relevant for network formation.

1.4.3 Adding in Scalar Unobservables

Recall that the purpose of the network formation model is to recover the unobserved variables a_{is} for each i in school s, in order to control for network endogeneity in the peer effects model. Having derived results for the simple model, I now add these into the model. These must be estimated in order to control for the endogeneity of P_{is} and $\bar{P}_{is}$ in the peer-effects outcome equation from the prior section.

1.4.3.1 Equilibrium and Functional Form

Scalar unobservables a_{is} and a_{js} are included in the model as part of the function f.[23] Functionally, they enter utility exactly the same way as X_{is} and X_{js}. That is, these scalar unobservables change the relative utilities of the various linking strategies. I make the following assumption on the functional form of f:

$$\begin{aligned}\tilde{f}(X_{is}, X_{js}, a_{is}, a_{js}) = {} & \gamma_1 X_{is} + \gamma_2 a_{is} + \delta_1 X_{is} X_{js} + \delta_2 X_{is} a_{js} + \delta_3 a_{is} X_{js} \\ & + \delta_4 a_{is} a_{js} + \gamma_3 X_{js} + \gamma_4 a_{js} \end{aligned} \tag{1.26}$$

In all specifications, the vector of observed variables X_{is} contains a participation indicator P_{is}. In order for omitted a_{is} to bias estimates of the peer effects model, it must change the

[23] As such, the utility function shares some features with that employed by Graham (2014*a*). However, in contrast to that model, I allow a_{is} and a_{js} to interact with each other and also with observable characteristics X_{is} and X_{js}.

relative utilities derived from links conditional on X (and thus P). Note the centrality of the *interactions* between a and X here. For example, if δ_3 is positive, then individuals with higher a_{is} derive more utility from linking with participant individuals (for whom $P_{js} = 1$) than those without such characteristic (for whom $P_{js} = 0$). This leads them to have higher $\bar{P}_{is}$ in the outcome equation, which is clearly correlated positively with a_{is}.

With the additional assumption that a_{is} is bounded, the equilibrium results for the simple case extend to the case with scalar unobservables. That is, the results in Propositions 1.2 through 1.5 hold. Equilibria exist, the strictly positive equilibrium is unique, and the results for potential games follow as well.

1.4.3.2 Identification Results with Scalar Unobservables

The simple model effectively assumes $a_{is} = 0$ for every individual. This rules out the primary source of endogenous network formation that leads to bias in the peer effects estimates. Adding these back into the model, Equation (1.20) becomes Equation (1.27).

$$\dot{g}^i_{ijs} = \tilde{\beta}\dot{g}^i_{jis} + \delta_1 X_{is}\dot{X}^i_{js} + \delta_2 X_{is}\dot{a}^i_{js} + \delta_3 a_{is}\dot{X}^i_{js} + \delta_4 a_{is}\dot{a}^i_{js} + \gamma_3\dot{X}^j_{is} + \gamma_4 a_{is}\dot{a}^i_{js} - \dot{c}^i_{ijs} \tag{1.27}$$

Again, the mirror image of Equation (1.27) provides instruments for endogenous $\dot{g}^i_{jis}$.

$$\dot{g}^j_{jis} = \tilde{\beta}\dot{g}^j_{ijs} + \delta_1 X_{js}\dot{X}^j_{is} + \delta_2 X_{js}\dot{a}^j_{is} + \delta_3 a_{js}\dot{X}^j_{is} + \delta_4 a_{js}\dot{a}^j_{is} + \gamma_3\dot{X}^j_{js} + \gamma_4 a_{js}\dot{a}^j_{is} - \dot{c}^j_{jis} \tag{1.28}$$

Assumption 1.6 provides exogeneity assumptions for the full model.

Assumption 1.6. *The following exogeneity conditions hold:*

1. $\mathbb{E}[\log c_{ijs}|\mathbf{X_{ks}}, a_{ks}] = 0 \;\forall\; k$
2. $\mathbb{E}[a_{is}|\mathbf{X_{ks}}] = 0 \;\forall\; k$
3. $\mathbb{E}[a_{is}|a_{js}] = 0 \;\forall\; j \neq i$

The first part of the assumption is similar to Assumption 1.4 and implies that unobserved costs are (mean) independent of individual-level observed and unobserved variables. The second part serves to separate the composite term $(\gamma_3 X_{js} + \gamma_4 a_{js})$,[24] while the third part of Assumption 1.6 rules out correlation among these unobserved variables.

Assumption 1.7. *$(\tilde{\beta}, \delta_1', \delta_2', \delta_3', \delta_4, \gamma_3', \gamma_4) \in \mathbf{\Theta}$, a compact set in $\mathbb{R}^{4m+3}$, where $m = dim(X_{is})$. Further, $a_{js} \in \mathbf{\Omega}\ \forall\ j, s$, where $\mathbf{\Omega}$ is a compact set in $\mathbb{R}$.*

Identification results are analogous to those of the simpler model. Proposition 1.7 states the first result.

Proposition 1.7. *Define $z_{ijs} = [X_{is}\dot{X}^i_{js}, \dot{X}^i_{js}, \frac{1}{N_s-2}\sum_{k\neq i,j}[X_{ks}, X_{is}X_{ks}, X_{js}X_{ks}]]$ and $b_{ijs} = [\dot{g}^i_{jis}, X_{is}\dot{X}^i_{js}, \dot{X}^i_{js}]$. Given Assumptions 1.3, 1.6, and 1.7, the parameters $\tilde{\beta}$, γ_1, and δ_1 are identified if $\mathbb{E}[z_{ijs}' b_{ijs}]$ is of rank $l \geq 2m+1$, where $m = dim(X_{is})$.*

Proof. See Appendix A. □

Additional assumptions are necessary to identify the remaining network-formation parameters, as stated in Assumption 1.8. Proposition 1.8 provides conditions under which the parameters are identified, but only to scale. These parameters are only identified to scale due to the fact that we can re-scale them by correspondingly re-scaling the latent variable a. For a given normalization, such as $\sigma_a^2 = 1$, these parameters are identified absolutely.

Assumption 1.8. *The following conditional variance restrictions hold:*

1. $\mathbb{E}[a_{is}^2 | \mathbf{X_{ks}}] = \sigma_a^2\ \forall\ k$

2. $\mathbb{E}[a_{is}^2 | a_{js}] = \sigma_a^2\ \forall\ j \neq i$

Proposition 1.8. *Given Assumptions 1.6 and 1.8, the parameters γ_2, δ_2, δ_3, and δ_4 are identified to scale if the following rank conditions hold:*

[24] This assumption imposes independence between observed and unobserved variables, an assumption also made by Hsieh and Lee (2016).

1. $rank(\mathbb{E}[z'_{ijs}b_{ijs}]) = 2m+1$, *where* $b_{ijs} = [\dot{g}^i_{jis}, X_{is}\dot{X}^i_{js}, \dot{X}^i_{js}]$

2. $rank(\mathbb{E}[z'_{ijs}b^2_{ijs}]) = m+1$, *where* $b^2_{ijs} = [1, X_{js}]$

3. $rank(\mathbb{E}[z'_{ijs}X_{is}]) = m$

4. $rank(\mathbb{E}[z'_{ijs}]) = 1$

Proof. See Appendix A. □

Recall that the peer effects model relies upon controlling for latent variables a_{is}, and thus we need to recover these variables in order to identify its parameters. While results to this point have relied upon "many market" asymptotics, identification of a_{is} relies upon "large market" asymptotics. For an individual in a school of size N_s, we observe $N_s - 1$ links.

Proposition 1.9. *For a given* s, $\gamma_4 + \delta_2 \underset{i \neq j}{\mathbb{E}}[X_{is}] \neq 0 \Rightarrow a_{js}$ *is identified to scale for all* j *as* $N_s \to \infty$.

Proof. See Appendix A. □

Proposition 1.9 provides the main identification result for scalar unobservables a_{is}. The condition that $\gamma_4 + \delta_2 \underset{i \neq j}{\mathbb{E}}[X_{is}] \neq 0$ is a relevance condition requiring that unobserved a_{is} actually play a part in the network-formation process. If X_{is} includes *all* variables involved in determining network links, then this condition will fail, but this seems quite unlikely. Note that, as in Proposition 1.8, each a_{is} is only identified to scale, a scale that can be fixed with a convenient normalization.

These three propositions provide the primary identification results for the model with scalar unobservables. Observation of many networks provides identification of parameters common to all networks, as given in Propositions 1.7 and 1.8. Observation of a large number of linkages within each network provides identification of the vector of individual-specific parameters a_{is} for each i in each s, as given in Proposition 1.9. Note again that, similar to the simple case in the prior subsection, parameters that involve variables that vary only with

i are not identified. As in the simpler case, this non-identification result essentially amounts to the inability to make welfare claims about different network configurations, which are thus beyond the scope of this paper.

1.5 Structural Estimation Results

Armed with the identification results from the previous two sections, I now proceed to structural estimation. For the purposes of estimation, I restrict attention to Random Treatment and Control schools, setting aside Elected Treatment for use in the validation exercise in Section 1.6.

Structural estimation consists of two steps. First, I estimate the parameters of the network formation game. Next, conditional on these parameters—particularly the estimated structural unobservables a_{is}—I estimate the parameters of the outcome equation which accounts for peer effects. These estimates indicate that unobserved a_{is} plays a large role in the determination of both network structure and outcomes conditional on network structure. Further, I provide evidence that, at least in this case, failure to account for a_{is} leads to crucially biased estimates of peer effects parameters.

1.5.1 Network Formation Estimation

This section estimates the network formation model using the identification results in the prior section. First, I discuss how I handle missing network data and zeros in the network data. Then, I estimate the parameters of the network formation model. As expected, I see that the process exhibits substantial homophily, and I further show that the latent variables a_{is} play an important part in network formation.

1.5.1.1 Missing Network Data

As described in Section 1.2 above, network data was collected via school visits after the conclusion of the Bal Sabha program. Accordingly, we have missing network data for two reasons. First, some students were not present in school on the date of the survey. Second, students may not have properly answered the survey questions. Of the estimation sample in the two treatment arms used for structural estimation (Control and Random Treatment), link values for approximately 40% of possible link pairs is missing.

Missing network data has the potential to confound estimation for a number of reasons. If data is missing non-randomly, listwise deletion leads to biased estimates of even network-level descriptive statistics (see, e.g., Chandrasekhar and Lewis, 2011). In the specific model outlined here, missing network data means that we do not observe an individual's entire vector of network choices. If certain types of students, defined by observed or unobserved characteristics, are more likely to be absent on the day of the network survey, then we need a way of accounting for these students.

Accordingly, a method of reconstructing the missing network data is needed. Chandrasekhar and Jackson (2014), using a model arising from the random graphs literature, provide a method that reconstructs networks based upon the probability of observing given dyadic and triadic relationships in the data. Williams (2016) recently extended this method to allow for missingness to vary by observed characteristics. He shows that the method does a reasonable job in reconstructing missing data in AddHealth with 75% missing data, as in his application.[25] He then applies the method to simulate missing network data at the Air Force Academy. A key limitation of this method, however, is that it does not model tradeoffs between linking strategies.

Fortunately, in my context, the network formation model can be pressed into service to fill in missing data. In his application at the Air Force Academy, Williams (2016) does not model network formation; rather, he models only outcomes conditional on the observed network. In contrast, I have posited a specific model of network formation that can be used to reconstruct missing data.

Accordingly, I employ an iterative EM algorithm that uses the network model itself to

[25] While the network reconstruction technique I employ is different, I note that my network data has a much higher response rate (60% vs. 25%) than the data employed by Williams (2016).

simulate missing data. Details are in Appendix D, but the basic structure is described as follows. First, fill in the missing data arbitrarily. Second, estimate the network formation model with this dataset. Third, using these estimated parameters and implied distributions of unobserved data, simulate values for missing network data. Repeat the second and third steps for sufficient iterations to converge to the distribution of both the simulated networks and the estimated parameters. This generates a Markov Chain of simulated networks and estimated parameters.[26] After a sufficient burn-in period, I take draws from this chain as the simulated parameters and full networks.[27]

1.5.1.2 Zeros in the Data

Recall that identification of the parameters of the network-formation game depends on observing the *strictly positive* equilibrium. This implies that no pairs of students choose a zero link to each other. In the actual data, however, there are a number of students who answer all of the link questions negatively, leading to the constructed continuous link measure being zero. In the raw link data, of 58,530 dyads used for estimation, approximately 28.5% are zeros.

I attribute these zeros to measurement error. That is, the actually-observed continuous network measure is a noisy version of the true measure.[28] It is constructed from nine binary questions. Presumably, especially given potential networks that average 44 students, if we asked substantially more link questions, the answers to some questions would be positive. Accordingly, in order to account for this, whenever zeros appear in the constructed continuous link measure, I replace this value with an imputed value that is drawn randomly and uniformly between 0 and the minimum link measure observed in the actual (non-zero)

[26] While the raw data consists of discrete network measures, the network formation model operates at the level of continuous link values. Accordingly, the imputation algorithm—which employs the network formation model—directly imputes the continuous measure.

[27] In practice, after a burn-in period of 100 iterations, I take 20 draws with a gap of 10 between each draw.

[28] I note that measurement error is theoretically a issue for all network data, not just those that are observed as zero in the data. A more formal model might account for error in constructing the continuous network measure, for example, analogously to Cunha, Heckman and Shennach (2010) in their study of educational skills formation. Future projects with the continuous network link models will explore this issue further.

data.[29]

1.5.1.3 Network Formation Parameters

Having discussed the data issues, I now move on the estimation. Essentially, this consists of finding values of the structural parameter $(\tilde{\beta}, \gamma, \delta)$ and the scalar unobservables a such that the sample analogues of the assumed moment conditions hold empirically.[30] I estimate this via GMM, with moments motivated by the identification results.[31] As discussed in the prior subsection, in order to correct for missing data, the GMM routine is the minimization step of the iterative EM algorithm.

Estimated parameters of the network formation game are given in Table 1.10. In Panel A, we see that estimated $\tilde{\beta}$ is positive and highly significant, indicating that effort levels of the two actors forming a link are strongly complementary, consistent with the reduced-form facts. Importantly, this is true even when controlling for a large set of observed and unobserved characteristics. Further $\tilde{\beta}$, estimated at 0.207, is substantially less than one,[32] as required for the network-formation process to have a unique strictly positive equilibrium.

Additionally, the point estimate of γ_2 shows that scalar unobservables a are important in link decisions. The parameter γ_2 identifies the additional utility derived to individual i from linking with j when j's unobserved a_{js} increases by one standard deviation. Note that the effect of a one standard deviation change, 0.693, is of the same order of magnitude of the effect of homophily for many characteristics: for example, two students in Grade 7 derive 0.881 units more utility than if either is not in Grade 7.

[29] Estimates are not sensitive to simulation error. That is, the estimated network formation parameters are quite similar across many different draws of the algorithm. Further, estimates do not substantially differ between this imputation method and simply adding a small number, such as 0.001, to each observed link value.

[30] Recall that identification results showed that a_{is} and any parameters that interact with a_{is} are only identified to scale. I have set this scale by setting the variance of estimated a_{is} to one.

[31] An alternative and somewhat less computationally burdensome procedure estimates the structural parameters $\theta = (\tilde{\beta}, \gamma, \delta)$ and the scalar unobservables vector a_{is} iteratively. However, this procedure has the limitation that it does not allow for closed-form standard errors. Therefore, while point estimation is sped up, the need to bootstrap the procedure leads this procedure to be substantially slower overall due to increased burden of variance estimation.

[32] A one-sided test strongly rejects the null that $\tilde{\beta} \geq 1$.

Table 1.10: Structural Network Formation Parameter Estimates

Panel A: Parameters Not involving Covariates				
$\tilde{\beta}$	0.207***			
	(0.059)			
γ_2	0.693***			
	(0.023)			
δ_4	0.116***			
	(0.012)			
Panel B: Parameters involving Covariates				
X Variable	γ_1	δ_1	δ_2	δ_3
Elected	0.438***	0.052	-0.036	0.182***
	(0.026)	(0.043)	(0.023)	(0.022)
Grade 7	-0.103***	0.881***	0.022	0.001
	(0.022)	(0.069)	(0.019)	(0.019)
Grade 8	-0.134***	0.984***	-0.009	0.054***
	(0.023)	(0.079)	(0.019)	(0.019)
SC	-0.765***	0.939***	-0.245***	-0.059*
	(0.037)	(0.075)	(0.026)	(0.033)
ST	-0.765***	1.020***	-0.276***	-0.038
	(0.038)	(0.115)	(0.034)	(0.044)
OBC	-0.323***	0.229***	-0.072***	0.010
	(0.033)	(0.037)	(0.020)	(0.022)
Participant	-0.098	0.077	-0.426***	0.348***
	(0.118)	(0.077)	(0.129)	(0.123)
Participant × Elected	-0.131**	-0.108	0.061	-0.075
	(0.058)	(0.124)	(0.058)	(0.058)
Participant × Grade 7	-0.068	-0.345**	-0.010	-0.116*
	(0.068)	(0.169)	(0.069)	(0.066)
Participant × Grade 8	0.076	-0.282**	0.031	-0.072
	(0.073)	(0.116)	(0.062)	(0.059)
Participant × SC	0.466***	-0.500***	0.232*	-0.173
	(0.120)	(0.169)	(0.133)	(0.130)
Participant × ST	0.408***	-0.287**	0.272**	-0.306**
	(0.116)	(0.147)	(0.139)	(0.131)
Participant × OBC	0.055	0.211**	0.196	-0.341***
	(0.111)	(0.107)	(0.124)	(0.118)

Notes: N = 58,530 in 20 schools. Robust standard errors in parentheses, clustered by school. *** p<0.01, ** p<0.05, * p<0.1. Parameters estimated via GMM. Missing data imputed and estimates adjusted via algorithm described in Appendix D. SC = Scheduled Caste, ST = Scheduled Tribe, OBC = Other Backwards Caste. Omitted Categories are Not Elected, Grade 6, and General.

Panel B presents parameter estimates that show how observed and unobserved variables interact in determining the utility of network links. The γ_1 parameter identifies the difference in utility to individual i from linking with j when j has the indicated characteristic versus j not having it. For example, if j is elected, then i derives 0.438 units more utility than if j is not elected. I note that the negative point estimates on members of lower castes suggest less utility from linking with them but that these effects are mitigated somewhat when they are chosen to participate, as indicated by positive and significant coefficients for interactions of SC and ST with the participation indicator.

The second column indicates substantial homophily along a number of dimensions, as shown by the δ_1 estimates. Those in Grades 7 and 8 derive more utility from linking with their classmates, and members of Scheduled Castes, Scheduled Tribes, and Other Backwards Castes similarly get more utility from linking to others in the same population grouping. Interestingly, among those in Grades 7 and 8 as well as Scheduled Tribes and Scheduled Castes, being chosen to participate seems to mitigate homophilic tendencies, as indicated by the negative and significant point estimates of interactions between Participant and these characteristics.

The final two columns show estimates of the effects of interactions between observed characteristics and unobserved a. Many of these estimated coefficients are highly significant and large in magnitude. This suggests that these interactions are quite important in individuals' decisions about network formation. Accordingly, failure to account for these interactions has the potential to crucially bias estimates of the parameters of the peer effects model.

1.5.2 Peer Effects Estimates

This section presents estimates of the peer effects model, as specified by Equation (1.7). Similar to the network formation case, I first describe how I treat missing outcome data. Then I present the estimated parameters, which show that structural unobservables a are important in determining outcomes and that failure to account for them leads to biased estimates of parameters of the peer effects model.

1.5.2.1 Missing Outcome Data

Similar to the network link variables, I encounter missing data for two reasons. First, some girls were not present on the day that the endline questionnaires were administered. Second, even if they were present, some students did not answer the relevant questions.

To account for this possibly non-random missing data, I employ an iterative EM algorithm. Estimation is done by OLS, which then imputes outcomes according to the estimated distribution of unobserved variables. Importantly, the parameters of the peer effects model are estimated conditional on a realized network and unobserved parameters a_{is}. Accordingly, to account for variance in imputing the network data, I take draws from the imputed distribution of networks and unobserved a_{is}, as these were calculated as part of the network formation estimation process. Conditional on each draw of the network and a_{is},[33] I iterate the algorithm 500 times to minimize sensitivity to starting values.

1.5.2.2 Peer Effects Parameter Estimates

Tables 1.11 and 1.12 present estimated parameters of the peer effects model. These estimates are calculated via OLS conditional on the realized network and estimated a_{is}. From these, I construct $\overline{\text{Participant}}$ and $\bar{a}$. From this we see that latent variable a_{is} plays a large role in determining outcomes.

First, Table 1.11 gives results for Educational Aspirations. Column (1) estimates the simple model that does not account for baseline outcomes or a. Column (2) adds a and peer effects on a ($\bar{a}$). Columns (3) and (4) are analogous to Columns (1) and (2) but further include baseline outcomes and peer effects for baseline outcomes, both interacterd with $\hat{L}$, $\hat{M}$, and $\hat{H}$.

Since they are more general models, I focus discussion on Columns (3) and (4). In

[33] I take 20 such draws from the Markov Chain of the network data, from which I construct 20 imputations of the outcome data.

Column (4), we see positive and significant coefficients for the interaction between a and all three tercile indicators. Since a is normalized to have standard deviation of one, this means that, among those in the lowest predicted outcome tercile, one standard deviation higher unobserved a leads to 0.059 standard deviations higher predicted Educational Aspirations. We see similar positive effects for those in the other predicted terciles (0.070 and 0.095 for $\hat{M}$ and $\hat{H}$, respectively). Recall that, from the network formation estimates, we saw that individuals derive more utility from linking to those with higher a. Accordingly, this suggests that those who are more desirable as friends also have higher unobserved factors that affect their Educational Aspirations. While noting less power,[34] I note that there no significant coefficients on the interactions between $\hat{a}$ and predicted tercile indicators.

In addition to showing that the omitted a variables seem are important in determining outcomes in Column (4), I also test cross-equation restrictions between Equations (3) and (4). These test the equality of each set of interactions with $\hat{L}$, $\hat{M}$, and $\hat{H}$: for example, the first tests whether the three coefficients on Participant $\times\,\hat{\mathrm{L}}$, Participant $\times\,\hat{\mathrm{M}}$, and Participant $\times\,\hat{\mathrm{H}}$ are equal between Columns (3) and (4). I perform similar tests for $\overline{\text{Participant}}$, Baseline Outcome, and $\overline{\text{Basline Outcome}}$. P-values of each of these tests are presented at the bottom of Table 1.11, between Columns (3) and (4).[35] Each of these tests strongly rejects the null of equality, suggesting that the estimates in Column (3), which ignore network endogeneity, are biased.

Results for Gender Roles attitudes are presented in Table 1.12, which is structured similarly to Table 1.11. Again, I focus discussion on Columns (3) and (4). In Column (4), we see that the effect of unobserved a is significant for those in all three predicted terciles. However, the effect on those in the lowest tercile is negative: among students with the lowest predicted outcomes, one standard deviation higher a implies 0.071 standard deviations *lower* Gender Roles attitudes, all else equal. This is in contrast to the positive and significant effect of

[34] Standard errors are much higher for the $\bar{a}$ variables than the a variables. For example, compare the standard errors on the coefficients in Column (4) on $\overline{\mathrm{a}} \times \hat{\mathrm{L}}$ (0.033) versus $\mathrm{a} \times \hat{\mathrm{L}}$ (0.208). This is likely due to the fact that $\bar{a}$ is constructed from many estimates of a, all of which are noisily estimated.

[35] Similar tests are conducted on the relevant coefficients in Columns (1) and (2).

Table 1.11: Structural Peer Effects Estimates (Education)

	(1)	(2)	(3)	(4)
Participant × $\hat{\text{L}}$	-0.284***	-0.307***	-0.356***	-0.242***
	(0.031)	(0.044)	(0.030)	(0.042)
Participant × $\hat{\text{M}}$	0.012	-0.084***	-0.007	0.016
	(0.028)	(0.032)	(0.028)	(0.031)
Participant × $\hat{\text{H}}$	0.208***	0.162***	0.160***	0.252***
	(0.035)	(0.042)	(0.034)	(0.038)
$\overline{\text{Participant}}$ × $\hat{\text{L}}$	-0.524***	-0.418***	-0.371***	-0.915***
	(0.052)	(0.065)	(0.052)	(0.063)
$\overline{\text{Participant}}$ × $\hat{\text{M}}$	-0.060	0.160**	0.207***	0.158**
	(0.053)	(0.067)	(0.053)	(0.062)
$\overline{\text{Participant}}$ × $\hat{\text{H}}$	-0.267***	-0.380***	0.019	-0.095
	(0.088)	(0.108)	(0.081)	(0.094)
a × $\hat{\text{L}}$		0.005		0.059**
		(0.037)		(0.033)
a × $\hat{\text{M}}$		0.089***		0.070**
		(0.032)		(0.031)
a × $\hat{\text{H}}$		0.168***		0.095***
		(0.037)		(0.034)
ā × $\hat{\text{L}}$		-0.048		0.123
		(0.327)		(0.208)
ā × $\hat{\text{M}}$		-0.147		-0.129
		(0.121)		(0.113)
ā × $\hat{\text{H}}$		-0.027		0.022
		(0.155)		(0.150)
Baseline Outcome Interactions	NO	NO	YES	YES
$\overline{\text{Baseline Outcome}}$ Interactions	NO	NO	YES	YES
P-value of cross-equation tests for interactions with:				
Participant	0.000		0.000	
$\overline{\text{Participant}}$	0.000		0.000	
Baseline Outcome			0.000	
$\overline{\text{Baseline Outcome}}$			0.000	

Notes: Coefficients for $\hat{\text{L}}$, $\hat{\text{M}}$, and $\hat{\text{H}}$ suppressed. Robust standard errors in parentheses, clustered by school. *** p<0.01, ** p<0.05, * p<0.1. N = 920 in 20 schools in all specifications. Missing data imputed and estimates adjusted via algorithm described in Appendix D. Standard error calculations account for variance in estimating generated regressors a and ā. Baseline Outcome Interactions include interactions of Baseline Outcome with $\hat{\text{L}}$, $\hat{\text{M}}$, and $\hat{\text{H}}$. $\overline{\text{Baseline Outcome}}$ Interactions include interactions of $\overline{\text{Baseline Outcome}}$ with $\hat{\text{L}}$, $\hat{\text{M}}$, and $\hat{\text{H}}$. P-values at bottom of table represent tests of equality of interactions with the given variable and $\hat{\text{L}}$, $\hat{\text{M}}$, $\hat{\text{H}}$ across pairs of columns: e.g., the first p-value tests the equality of the three interactions with Participant across columns (1) and (2). All p-values calculated from $\chi^2(3)$

unobserved a for those in the middle and highest predicted terciles. There are no significant effects of $\overline{a}$, while again noting lack of power. Finally, I note that tests of equality of coefficients between Columns (3) and (4) strongly reject a null of equality, suggesting that estimates that fail to account for unobserved a are biased.

1.6 Out-of-Sample Validation

While a structural model allows for out-of-sample prediction, our confidence in the model can be bolstered by comparison of the model's predictions to realized out-of-sample outcomes. Fortunately here, I have an out-of-sample treatment group that can be used for this validation step, as suggested by Todd and Wolpin (2006). In Elected Treatment schools, which were not used in structural estimation in the prior section, participation in the program was assigned by election rather than randomly, as was done in Random Treatment schools. Therefore, having used Random Treatment and Control to estimate the model, I now use the estimated parameters to predict outcomes conditional on all participants being chosen by election. Comparing these predictions to the actual realized outcomes in Elected Treatment schools provides a check on the model's predictive power.

1.6.1 Simulation Method

Counterfactual simulation relies upon simulation of unobserved variables. The network-formation model includes three such unobservables: c_{ijs}, M_{is}, and a_{is}. The cost variables c_{ijs} are by construction independent of all observables and a_{is}. Accordingly, they are drawn from an independent log-normal distribution with mean zero and variance $\hat{\sigma}_c^2$, where $\hat{\sigma}_c^2$ is the empirical variance of these residuals from the estimation routine. Scalar variable a_{is} is similarly drawn from an independent normal distribution with mean zero and variance 1 (recall that this mean and variance are imposed as moment conditions). Finally, M_{is} is drawn from a

Table 1.12: Structural Peer Effects Estimates (Gender Roles)

	(1)	(2)	(3)	(4)
Participant × $\hat{\text{L}}$	0.065***	0.082***	0.108***	0.240***
	(0.024)	(0.028)	(0.024)	(0.027)
Participant × $\hat{\text{M}}$	-0.433***	-0.509***	-0.490***	-0.466***
	(0.029)	(0.034)	(0.029)	(0.033)
Participant × $\hat{\text{H}}$	0.308***	0.173***	0.322***	0.332***
	(0.035)	(0.047)	(0.034)	(0.041)
$\overline{\text{Participant}}$ × $\hat{\text{L}}$	-0.404***	-0.333***	-0.152***	0.070
	(0.044)	(0.058)	(0.046)	(0.056)
$\overline{\text{Participant}}$ × $\hat{\text{M}}$	-0.632***	-0.372***	-0.856***	-0.428***
	(0.051)	(0.071)	(0.049)	(0.064)
$\overline{\text{Participant}}$ × $\hat{\text{H}}$	-0.721***	-0.613***	-0.504***	-0.651***
	(0.061)	(0.092)	(0.058)	(0.078)
a × $\hat{\text{L}}$		-0.076**		-0.071**
		(0.030)		(0.030)
a × $\hat{\text{M}}$		0.038		0.069**
		(0.028)		(0.028)
a × $\hat{\text{H}}$		0.130***		0.082**
		(0.037)		(0.035)
ā × $\hat{\text{L}}$		-0.000		0.045
		(0.166)		(0.166)
ā × $\hat{\text{M}}$		0.077		-0.009
		(0.126)		(0.123)
ā × $\hat{\text{H}}$		0.187		0.054
		(0.205)		(0.195)
Baseline Outcome Interactions	NO	NO	YES	YES
$\overline{\text{Baseline Outcome}}$ Interactions	NO	NO	YES	YES
P-value of cross-equation tests for interactions with:				
Participant	0.000		0.000	
$\overline{\text{Participant}}$	0.000		0.000	
Baseline Outcome			0.000	
$\overline{\text{Baseline Outcome}}$			0.018	

Notes: Coefficients for $\hat{\text{L}}$, $\hat{\text{M}}$, and $\hat{\text{H}}$ suppressed. Robust standard errors in parentheses, clustered by school. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. N = 920 in 20 schools in all specifications. Missing data imputed and estimates adjusted via algorithm described in Appendix D. Standard error calculations account for variance in estimating generated regressors a and ā. Baseline Outcome Interactions include interactions of Baseline Outcome with $\hat{\text{L}}$, $\hat{\text{M}}$, and $\hat{\text{H}}$. $\overline{\text{Baseline Outcome}}$ Interactions include interactions of $\overline{\text{Baseline Outcome}}$ with $\hat{\text{L}}$, $\hat{\text{M}}$, and $\hat{\text{H}}$. P-values at bottom of table represent tests of equality of interactions with the given variable and $\hat{\text{L}}$, $\hat{\text{M}}$, $\hat{\text{H}}$ across pairs of columns: e.g., the first p-value tests the equality of the three interactions with Participant across columns (1) and (2). All p-values calculated from $\chi^2(3)$

log-normal distribution allowing for some dependence on observed characteristics.[36] I take the distribution of observed characteristics in Elected Treatment as given in all simulations in order to avoid any possible composition issues.

After the network-formation process is simulated, I move on to simulating outcomes. The parameters in Tables 1.11 and 1.12 are used to predict outcomes conditional on the simulated network and simulated a_{is}. Again, to avoid any composition bias, all simulations are done on 10 schools with the exact distribution of observed covariates as found in Elected Treatment schools.

In order to facilitate comparisons, I simulate outcomes under two different specifications. First, I generate simulations using "naive" estimates that ignore network endogeneity through a. The parameters of this model are drawn from Column (3) of Tables 1.11 and 1.12. Then, I generate simulations using the full model, corresponding to Column (4) of those same tables.

1.6.2 Comparison to Elected Treatment

Simulation results are presented in Table 1.13.[37] Simulations for Educational Aspirations are presented in Panel A, which shows that both simulation specifications are overly optimistic about mean Educational Aspirations. This could be due to a number of issues, including a possible discouragement effect of choice by election for those not selected. That is, it may be the case that the program carried out with elected participants serves to reinforce marginalization for girls not elected to the program. Still, while neither model can be rejected due to lack of power, the overall mean for the model that uses a is closer to the actual realized mean.

While the model's predicted Educational Aspirations may be biased upwards, it does a

[36] In practice, simulating M_{is} is a three-step process as follows: (1) for each i and s, recover $\hat{M}_{is}$ from the estimation routine, where $\hat{M}_{is} = \sum_{j \neq i} \hat{c}_{ijs} g_{ijs}$, (2) regress $\log \hat{M}_{is}$ on the same observed variables that appear in Table 1.10, (3) with these parameter estimates and implied variance of residuals $\hat{\sigma}_M^2$, simulate $\log M_{is}$, drawing the residuals from the a normal distribution with variance $\hat{\sigma}_M^2$.

[37] This exercise is analogous to that undertaken in Tables 12-15 in Todd and Wolpin (2006).

good job of getting at treatment effect heterogeneity. Consistent with the realized outcomes, we see that elected girls do better than those who were not elected, as do members of General castes. Members of lower castes are predicted to have substantially lower Educational Outcomes at endline, which matches the patterns from Elected Treatment.

The model does a better job of predicting the overall mean for Gender Roles attitudes, as shown in Panel B. Further, in contrast to the naive estimates (without a), the full model predicts substantially *lower* Gender Roles attitudes for non-elected girls than for their elected counterparts, a pattern that matches heterogeneity based on elected status in the realized data. Unfortunately, the model falls short in matching the pattern of heterogeneity across caste groupings.

In sum, the model does a credible job of matching many features of the out-of-sample realized outcomes. However, while I credibly match means within the entire sample for both outcomes, the model only predicts the patterns of heterogeneity for Educational Aspirations. This is qualitatively similar to the out-of-sample fit results in Todd and Wolpin (2006), who find that their model predicts average school attendance reasonably well for some subgroups but not others.[38]

1.7 Counterfactual Policy Evaluation

This section presents the results of simulated outcomes under counterfactual assignment rules. First, I present the results of simulations that assign individuals based on observed characteristics. Next, I present the effects of a policy that assigns based on unobserved (but estimated) a_{is}.

[38] A formal test for model fit as well as development of the theory of statistical power for such a test is beyond the scope of this paper but will be investigated in future work.

Table 1.13: Comparison of Realized to Predicted in Elected Treatment Schools

Panel A: Educational Aspirations							
	All	Elected	Not Elected	SC	ST	OBC	General
Observed in Elected Treatment							
Mean	-0.279	-0.053	-0.398	-0.364	-0.489	-0.392	0.085
Standard Error of Mean	0.151	0.180	0.163	0.233	0.179	0.189	0.234
N	330	114	216	64	32	153	81
Simulated with Same Covariate Distribution as Elected Treatment							
"Naive" Model (Without Scalar Unobservables a)							
Mean of Simulated Means	-0.165	-0.155	-0.171	-0.305	-0.295	-0.253	0.162
Full Model (With Scalar Unobservables a)							
Mean of Simulated Means	-0.198	-0.141	-0.228	-0.311	-0.310	-0.282	0.093

Panel B: Gender Roles Attitudes							
	All	Elected	Not Elected	SC	ST	OBC	General
Observed in Elected Treatment							
Mean	-0.022	-0.085	0.012	0.066	0.085	-0.339	0.462
Standard Error of Mean	0.147	0.210	0.149	0.240	0.239	0.121	0.100
N	332	116	216	65	33	153	81
Simulated with Same Covariate Distribution as Elected Treatment							
"Naive" Model (Without Scalar Unobservables a)							
Mean of Simulated Means	0.020	0.008	0.026	-0.051	-0.086	0.009	0.140
Full Model (With Scalar Unobservables a)							
Mean of Simulated Means	-0.034	-0.091	-0.004	-0.134	-0.098	0.016	-0.024

Notes: Standard errors clustered by school. Simulation results in Panel A correspond to Columns (3) (without scalar unobservables) and (4) (with scalar unobservables) of Table 1.11. Simulation results in Panel B correspond to Columns (3) (without scalar unobservables) and (4) (with scalar unobservables) of Table 1.12. Simulations based upon 10,000 repetitions, with residuals drawn from random normal. SC = Scheduled Caste, ST = Scheduled Tribe, OBC = Other Backwards Caste.

1.7.1 Assignment by Observed Variables

From the descriptive regression results in Table 1.2, we see that those in lower castes have substantially lower baseline outcomes, and those who were elected to participate have higher outcomes. Table 1.7, which was used to predict endline outcomes in the absence of treatment, shows similar patterns. Optimizing outcomes via creative assignment essentially amounts to choosing a set of individuals who will benefit the most from treatment, while accounting for the indirect effects of the program through network change and peer effects. Accordingly, an obvious place to start in thinking about alternative policies is to treat those who are in the most need.

Accordingly, I have designed three assignment rules.[39] Policy 1 assigns girls with the lowest predicted outcomes, as calculated by the estimates in Table 1.7. Policy 2 takes the opposite approach and selects those with the *highest* predicted outcomes. Finally, as a baseline for comparison, Policy 3 assigns girls to participate randomly, which is the same assignment rule actually implemented as part of the randomized evaluation in the Random Treatment schools.[40] I simulate outcomes in a similar manner to the algorithm described in the prior section, fixing the covariate distribution as that observed in Random Treament schools.

Table 1.14 demonstrates the results of these simulations. Note that the simulations were performed separately for the two outcomes, as the assignment rule does not in general assign the same girls under both outcomes. From this, we see that the assignemnt policy does not appear to have much effect on Educational Aspirations, but there are important distributional effects of the policies for Gender Roles.

First, looking at Educational Aspirations, the mean outcome is very similar across all

[39] Note that these assignment rules are all conditional on observed variables and not unobserved a. As a policy tool, assignment based on a—which is only recovered by estimation *after* the program outcomes are realized—seems implausible.

[40] Rather than simply use the realized outcomes in the data in the 10 Random Treatment schools, I simulate 1000 realizations under the same assignment rule. This allows for comparison to the *predicted* outcomes under random assignment rather than comparison to the single set of realized outcomes observed in the data.

three policies. Further, the means are also stable across the different subgroups. This indicates that the assignment rule has very little effect on mean Educational Aspirations.

In contrast, the assignment rule does appear to affect Gender Roles attitudes, as shown in Panel B. First, assigning those with the lowest predicted outcomes (Policy 1) leads to the highest mean outcome, and it also leads to much higher outcomes for members of Scheduled Tribes and Other Backwards Castes. Additionally, members of General castes perform substantially worse under the rule that preferentially assigns those with lowest predicted outcomes, especially as compared to Policy 2.

1.7.2 Assignment by Unobserved Structural Parameters

While less policy relevant, we can alternatively assign individuals by unobserved—but estimated—a_{is}. In the case of a program that runs over a single period, this exercise may be more academic. That is, in contrast to observed variables X_{is}, we do not observe a_{is} until after the program has run in a given school. However, in a setting where the program is run over multiple periods, we may estimate the model, including a_{is} in the first period then use the estimated a_{is} in evaluating potential assignment rules in later periods.

After we estimate a_{is}, we can see whether a policy that targets assignment by a_{is} would change outcomes. Similar to the prior section, I have designed two assignment rules. Policy 1 assigns those with the *lowest* estimated a_{is}, while Policy 2 assigns those wtih the *highest* estimated a_{is}. Simulations of Policy 3, assigning girls randomly, is included for comparison.

From this exercise, we see that assigning treatment to those with the *lowest* a_{is} improves both simulated average outcomes as compared to random assignment. Targeting to those wtih the *highest* estimated a_{is}, in contrast, brings about lower mean outcomes than random assignment.

Table 1.14: Counterfactual Policy Simulations (Assignment by Observed Variables)

Panel A: Educational Aspirations							
	All	Elected	Not Elected	SC	ST	OBC	General
Policy 1: Assign Girls with Lowest Predicted Outcome							
Mean of Simulated Means	-0.153	-0.157	-0.151	-0.337	-0.287	-0.090	0.285
S.D. of Simulated Means	0.056	0.092	0.065	0.102	0.133	0.075	0.146
Policy 2: Assign Girls with Highest Predicted Outcome							
Mean of Simulated Means	-0.152	-0.164	-0.147	-0.330	-0.273	-0.082	0.217
S.D. of Simulated Means	0.050	0.087	0.058	0.092	0.114	0.071	0.145
Policy 3: Assign Girls Randomly							
Mean of Simulated Means	-0.147	-0.154	-0.144	-0.329	-0.275	-0.083	0.274
S.D. of Simulated Means	0.053	0.091	0.061	0.098	0.125	0.073	0.145
Panel B: Gender Roles Attitudes							
	All	Elected	Not Elected	SC	ST	OBC	General
Policy 1: Assign Girls with Lowest Predicted Outcome							
Mean of Simulated Means	-0.021	-0.083	0.005	-0.123	0.034	0.033	-0.093
S.D. of Simulated Means	0.046	0.085	0.055	0.089	0.110	0.068	0.145
Policy 2: Assign Girls with Highest Predicted Outcome							
Mean of Simulated Means	-0.033	-0.083	-0.012	-0.083	-0.068	-0.021	0.112
S.D. of Simulated Means	0.046	0.085	0.054	0.088	0.111	0.068	0.144
Policy 3: Assign Girls Randomly							
Mean of Simulated Means	-0.056	-0.116	-0.031	-0.136	-0.041	-0.021	-0.028
S.D. of Simulated Means	0.048	0.086	0.056	0.091	0.111	0.069	0.145

Simulation results in Panel A correspond to Column (4) of Table 1.11. Simulation results in Panel B correspond to Column (4) of Table 1.12. Simulations based upon 1000 repetitions, with residuals drawn from random normal. SC = Scheduled Caste, ST = Scheduled Tribe, OBC = Other Backwards Caste.

Table 1.15: Counterfactual Policy Simulations (Assignment by Estimated Unobservable)

Panel A: Educational Aspirations

	All	Elected	Not Elected	SC	ST	OBC	General
Policy 1: Assign Girls with Lowest Estimated Unobservable							
Mean of Simulated Means	-0.113	-0.134	-0.105	-0.280	-0.238	-0.053	0.278
S.D. of Simulated Means	0.051	0.087	0.059	0.092	0.117	0.071	0.145
Policy 2: Assign Girls with Highest Estimated Unobservable							
Mean of Simulated Means	-0.186	-0.194	-0.183	-0.363	-0.316	-0.129	0.250
S.D. of Simulated Means	0.055	0.093	0.063	0.099	0.128	0.074	0.147
Policy 3: Assign Girls Randomly							
Mean of Simulated Means	-0.147	-0.161	-0.140	-0.318	-0.272	-0.088	0.263
S.D. of Simulated Means	0.052	0.088	0.060	0.096	0.121	0.072	0.142

Panel B: Gender Roles Attitudes

	All	Elected	Not Elected	SC	ST	OBC	General
Policy 1: Assign Girls with Lowest Estimated Unobservable							
Mean of Simulated Means	-0.035	-0.090	-0.012	-0.124	-0.014	0.007	-0.025
S.D. of Simulated Means	0.047	0.084	0.054	0.089	0.113	0.068	0.141
Policy 2: Assign Girls with Highest Estimated Unobservable							
Mean of Simulated Means	-0.077	-0.124	-0.058	-0.129	-0.075	-0.047	-0.079
S.D. of Simulated Means	0.047	0.085	0.055	0.089	0.110	0.067	0.146
Policy 3: Assign Girls Randomly							
Mean of Simulated Means	-0.052	-0.107	-0.030	-0.128	-0.042	-0.016	-0.033
S.D. of Simulated Means	0.047	0.085	0.055	0.090	0.115	0.068	0.143

Simulation results in Panel A correspond to Column (4) of Table 1.11. Simulation results in Panel B correspond to Column (4) of Table 1.12. Simulations based upon 1000 repetitions, with residuals drawn from random normal. SC = Scheduled Caste, ST = Scheduled Tribe, OBC = Other Backwards Caste.

1.7.3 Optimal Treatment Assignment

Assessment of the effects of counterfactual assignment policies in this context is an exercise in statistical treatment assignment (see Manski, 2004; Smith and Staghøj, 2009). That is, we search for statistical rules that maximize some function of outcomes, conditional on observable characteristics of individuals.[41] As pointed out by Bhattacharya (2009), the maximand under optimal assignment weakly dominates the maximand under any feasible assignment, including those used in the counterfactual simulations in the prior section. Accordingly, we need some way of assessing the effects of alternative assignments across the entire class of feasible alternative assignments. Dehejia (2005) formulates the problem in a Bayesian framework, drawing inferences from comparing features of the posterior predictive distributions. Similarly, Bhattacharya (2009) investigates the assignment of freshmen to dorms as a linear programming problem, providing results for maximizing the mean or any quantile of the outcome of interest.

The hypothetical optimal assignment problem here is complicated, however, by at least three factors. First, the presence of a budget constraint—only 13 girls per school can be assigned to program participation—complicates analysis. That is, even in settings in which agents' outcomes are independent, the need to estimate the threshold assignment rule—possibly including which covariates to include in this estimation—adds an important dimension of uncertainty that must be accounted for (Bhattacharya and Dupas, 2012).

Second, treatment externalities in the form of peer effects increase the complexity of the assignment problem. That is, identification and inference in the models from the econometrics literature typically rely upon independence across observations (see, e.g., Bhattacharya and Dupas, 2012; Manski, 2004). When, on the other hand, agents' outcomes are not independent, it may be impossible to derive a closed-form solution to the optimization problem. Accordingly, solving for the optimal assignment may necessitate the use of high-dimensional

[41] A slight twist on this formulation is provided in the prior section, where I predicted outcomes conditional on both observed variables and structural unobservables a_{is} that are recovered as part of the estimation procedure.

numerical programming procedures, an approach taken by Carrell, Sacerdote and West (2013).

Finally, and most pertinent to the central theme of this paper, the need to account for network endogeneity increases the complexity of the problem. In simulating the effects of alternative assignment using the model here, we need to simulate the network formation process to equilibrium at each potential assignment. Carrell, Sacerdote and West (2013) faced an already-high computational burden in solving for optimal assignment without taking account of this additional complication. Accordingly, using the model here to solve for optimal allocations entails an extreme computational burden and thus lies beyond the scope of the current project. Formulating methods to reduce this burden is a topic for future research.

1.8 Conclusion

The very existence of peer effects implies that individuals' outcomes and choices may be affected by the presence or absence of others. This suggests that, in settings where policymakers have control over assignments, simply changing the assignment rule may change outcomes, and such interventions should be relatively costless to implement. However, prior efforts to design and implement such assignment rules have fallen short due to endogenous peer selection. To account for this, we need a model of peer selection, but modeling and estimating such models presents many difficulties.

My approach explicitly models outcomes as the result of a two-step process. In the first step, agents choose peers within a continuous action space subject to a budget constraint. I show that this greatly simplifies equilibrium characterization and identification: under certain conditions, there is a unique strictly positive Nash equilibrium, and the first-order conditions implied by this can be employed for identification and estimation. The structure of the game motivates the use of a budget-set instrument to identify the model's parameters, and I provide conditions under which identification holds. Crucially, the model provides for identification of individual-specific unobserved variables that affect both network structure and outcomes.

In the second step, outcomes are determined conditional on the realized network. Here, network endogeneity is modeled explicitly as an omitted variable issue. Conditional on these unobserved variables—which are identified in the network formation model—the peer effects model is identified, even under certain types of network endogeneity.

With these methodological results in hand, I then estimate the model using innovative new data from a randomized trial in rural Rajasthan, India. I find that the unobserved variables play a large role in determining both network structure and outcomes conditional on the network. Further, statistical tests strongly reject a simpler model that ignores network endogeneity.

With the estimated parameters in hand, I move to out-of-sample validation and counterfactual simulation. First, by comparing predicted outcomes to realized out-of-sample means, I show that the model performs well in out-of-sample prediction. Next, I simulate the effects of alternative policies that assign participation preferentially. I separately assess the effects of policies that assign preferentially based upon observed variables and estimated unobservables.

With this paper, I provide a method to account for network endogeneity when estimating peer effects, giving an explicit model of how network endogeneity biases results that neglect to account for endogenous network structure. This further leads to a method to predict the effects of alternative assignments while accounting for network endogeneity. As necessary steps in developing this methodology, I make further contributions to the theory of network formation as well as providing new econometric results for the identification of network formation games.

Finally, this paper suggests a method that brings together structural and experimental methods of program evaluation. Importantly, the methodology developed here is not context-specific; rather, it has broad applicability in settings where assignment rules may influence outcomes both directly and through changing network structure. In order to apply the methods used here, researchers need to collect data on the outcome of interest and demographics, as well as sufficiently rich network data from which to construct a continuous network measure. With such data in hand, the researcher can then generate predictions of the effects of out-of-sample assignment rules that are robust to certain types of network endogeneity.

CHAPTER 2

How Many Friends Do You Have? An Empirical Investigation into Censoring-Induced Bias in Social Network Data

Abstract

In analyzing peer effects in a linear-in-means framework, identifying who interacts with whom is crucial. This suggests the need to collect detailed network data. However, taking a cue from AddHealth, many data-collection efforts only permit resondents to list up to a maximum number of links, leading to censoring and mismeasurement of peer groups. Within a linear-in-means framework, I document the extent of bias due to censoring analytically and by simulation. I then demonstrate that censoring-induced bias is present in empirical applications using data from AddHealth and an experiment in rural Nepal. After documenting the bias, I provide strategies to recover consistent estimates and discuss limitations of these strategies. This paper provides important contributions to the literature on design of network surveys as well as estimation of peer effects in the presence of data limitations.

2.1 Introduction

In economics and related fields, a multitude of papers have studied the effect of one's peers on one's own outcomes, especially in educational contexts. Before doing so, however, we must define the relevant peer group. For example, in school settings, many authors have defined the relevant peer group as those students who share a grade, classroom, or living space (See Sacerdote (2011) for a review of this literature).

More recently, economists have recognized that all others with whom an individual may potentially interact may not be the relevant peer group for purposes of peer effects. Rather, it is likely that an individual's actual social connections influence behavior and outcomes to a different degree than those with whom they share a classroom or school but rarely interact with. This has led to defining an individual's peer group as a subset of possible available peers.

Accordingly, in order to define peer groups more accurately, many studies now collect detailed network data. Such data often asks individuals to identify their friends, under the assumption that friends are the relevant influencers. A common feature of this data, however, is the presence of censoring. Censoring occurs when individuals may list only a certain number of network links. If for example, a survey only allows individuals to list up to five friends, and and individual has eight friends, then the final three friends are censored. This leads to mismeasurement of the relevant peer group.

The practice of gathering censored peer data dates back at least to AddHealth (Harris, 2009). In the Wave 1 survey, individuals are prompted to list their male and female friends, but only up to five of each. The censored friendship data from AddHealth has been used to study peer interactions in a wide number of applications, often with little sensitivity to the issue of censoring.

Moreover, the censoring issue is not limited to AddHealth. Rather, with a limited number of exceptions, many studies, especially in developing countries, allow individuals to list only a certain number of network links. Further, in many such studies, a large proportion of respondents list the maximum allowed number of links, suggesting that censoring may be a quantitatively significant issue. While many authors are aware of the potential of censoring-induced bias in peer effects estimates, there has been little effort made to quantify the potential bias or to devise strategies for dealing with this bias in censored data. This paper fills that gap in the literature.

In doing so, I make two primary contributions. First, I document the potential for censoring-induced bias in peer effects estimates. I do this analytically, by simulation, and finally by estimation using real datasets. I first derive analytic expressions for the bias of two commonly-used estimators of peer effects. Then I simulate data and estimates varying the number of friends observed, showing that when only few friends are observed, parameter estimates are meaningfully biased. As expected, as more friends are observed, censoring-induced bias disappears, and the estimates converge to the true paraemter values. Third, using data from two sources, I estimate parameters of peer effects models under different censoring rules, demonstrating that estimates are sensitive to the number of links observed.

Second, after demonstrating the bias in estimates due to censoring, I move on to the paper's second contribution: providing strategies to recover unbiased estimates. I provide two such strategies. The first employs an uncensored subset of networks to estimate an analytic bias correction. I show that this works well in simulations when a 10% uncensored subsample is collected. Second, I adapt an estimator from Chandrasekhar and Lewis (2011) to the setting at hand, which estimates the parameters of the model from those individuals for whom we observe all links. By simulation, I show that this strategy works well under restrictive assumptions.

Two papers in particular are relevant to the issues at hand. First, Chandrasekhar and Lewis (2011) deals with a related but different problem: networks sampled at the individual (dyad) level. This is a related but conceptually different problem than the one investigated here, where data is missing at the edge level and is not missing randomly. Their primary bias correction method relies upon graphical reconstruction, which I discuss briefly in Section

2.8. Second, Sojourner (2013) investigates the bias induced by missing peer covariate data in estimates of peer effects in the Tennessee STAR project. While likely relevant in many of the same contexts, the analysis deals with a classroom setting where data on individual interactions is unavailable.

This paper proceeds as follows. Section 2.2 discusses the extent of bias in network surveys, documenting that large proportions of respondents list the maximum number of links in many surveys. Section 2.3 presents the estimators under study and provides analytic expressions for censoring-induced bias. Section 2.4 then simulates data to show that these estimators are indeed biased when less than the full network is observed. Next, Section 2.5 draws upon two datasets to provide evidence that the simulated patterns of bias show up in real datasets. Section 2.6 then provides two strategies to correct for censoring-induced bias. Section 2.7 then shows that these strategies are likely to fail in the presence of substantial heterogeneity in parameters. Section 2.8 discusses the potential for graphical reconstruction, while Section 2.9 concludes.

2.2 Censoring in Network Surveys

2.2.1 Survey Questions that Induce Censoring

A widespread practice in collecting network data is to ask individual respondents to name their network links. This practice dates at least to AddHealth, which elicited network data with the following prompt:

> List your closest male friends. List your best male friend first, then your next best friend, and so on. Girls may include boys who are friends and boyfriends (Harris, 2009).

Each student in the sample is given this prompt as well as a prompt to list female friends. Crucially, students were allowed to list only up to five male and five female friends.

Prompts of this type are widespread, especially in surveys collected as part of field projects in developing countries. For instance, in their large-scale social data collection project in India, Banerjee et al. (2012) allow respondents to list up to either five or eight links along a number of dimensions. In a study on adoption of sanitary products by teenage girls in Nepal, Oster and Thornton (2012) allow respondents to list up to three close friends. In studying the interaction between social networks and insurance take-up, Cai, Janvry and Sadoulet (2015) allow respondents to list up to five close friends. Kandpal and Baylis (2013), in studying peer networks among women in Uttarakhand, India, similarly allow respondents to name up to five friends.

2.2.2 Exceptions to this Practice

While censoring is quite common in network data, it is far from universal. Two separate approaches seek to collect uncensored network data.

First, some have addressed the issue of censoring by collecting ordered connection data but with no upper bound on the number of links. For example, in a study of the effect of network links on learning HIV status in Malawi, Ngatia (2015) allows respondents to list as many social contacts as they desire, with some listing as many as 13 friends and 12 relatives. Similarly, Comola and Prina (2014), in their study of informal financial transactions, allow respondents to list relationships without limit.

Second, another line of surveys have dispensed with the practice of asking people to list their contacts. That is, these surveys have prompted individuals to provide their relationship to individuals identified by the survey instrument. In a study on adoption of new crop technology in Ghana, Conley and Udry (2012) ask respondents their relationships to randomly-chosen other individuals in their given village. Delavallade, Griffith and Thornton (2016) take a related approach but achieve full coverage rather than random sampling: each adolescent girl in their survey reveals her relationship to each other one in the school..

2.2.3 A Rough Measure of the Extent of Censoring

Clearly, censoring leads to mismeasurement of relevant networks under study. The extent to which this mismeasurement affects estimates of parameters of interest undoubtedly depends on the amount of data points that are censored. We can get a rough handle on the amount of potential censoring by looking at the number of individuals who list the maximum number of possible links. From reports in papers of the extent of censoring as well as looking at the datasets used in the empirical analysis here, it is apparent that there is substantial censoring in network datasets.

If a respondent lists the maximum number of links allowed, then there are two possibilities for his number of links. On the one hand, his number of links may be equal to the number allowed on the survey, in which case censoring does not lead to measurement error in his network. Alternatively, he may have more links than he is permitted to list, in which case the censoring rule leaves his peer group mismeasured. Accordingly, the number of respondents who list the maximum allowable number of links provides an upper bound on the extent of censoring in a given dataset.

Since it represents an upper bound, we can get a rough idea of the extent of censoring by looking at the number of respondents who list the maximum allowable number of links. At one of of the spectrum, Banerjee et al. (2012) show that less than 0.1% of their survey respondents name the maximum number of links (either 5 or 8) along any of the many dimensions they survey. In this setting, censoring is unlikely to cause much of an issue, although partial sampling is still an issue in that dataset (Chandrasekhar and Lewis, 2011). Alternatively, a large proportion of respondents naming the maximum number of links is a prominent feature of many network studies. In their study on the interaction between social networks and insurance take-up, Cai, Janvry and Sadoulet (2015) report that the majority of their respondents list five links, with an average number of 4.9.

Further, a large fraction of survey respondents in the two datasets I use in the empirical analysis below name the maximum number of possible links. Table 2.1 presents the percentage of respondents of each gender who name *at least* the number of nominees of the indicated genders. For example, 66.1% of female respondents name five female friends, while only 37.4% of female respondents name at least five male friends. From the final row, we see that a large percentage of respondents in AddHealth name the maximum number of friends allowed by the survey instrument.

Table 2.1: Friendship Nominations in AddHealth

Nominator	Female		Male	
Nominee	Female	Male	Female	Male
1	0.862	0.635	0.770	0.778
2	0.842	0.578	0.736	0.750
3	0.799	0.506	0.660	0.694
4	0.735	0.433	0.571	0.623
5	0.661	0.374	0.499	0.561

Table presents probability that Nominator of specified gender names *at least* that number of Nominees (friends) of the specified gender.

The data collected by Oster and Thornton (2012) follows a similar pattern. In that dataset, 68% of sampled girls report the maximum allowed three friends. Accordingly, in these two datasets, there is substantial potential for the data collection method, which allowed for censoring of peer groups, to bias estimates of parameters of interest. The remainder of this paper investigates the implications of this possible censoring for estimates of peer effects parameters.

2.3 Characterizing Bias

This section derives expressions for inconsistency induced by censoring. First, I introduce the estimators under study. Then, I derive analytic expressions for censoring-induced bias.

2.3.1 Data-Generating Process

The data-generating process takes place in two steps. First, networks are formed. Second, conditional on the realized network, outcomes are determined in a way that allows for peer effects.

2.3.1.1 Network Formation

The first step in determining outcomes is network formation. I make minimal assumptions on the process that generates networks. Links either exist or do not exist.The existence of a link between individuals i and j in school s is a binary variable, so $l_{ijs} \in \{0,1\}$.[1] For simplicity, I further assume that links are symmetric, so that $l_{ijs} = l_{jis}$.[2]

For each school s, there exists a matrix of links $\mathbf{L_s}$. In a school with K individuals, $\mathbf{L_s}$ is $M \times M$. Elements along the diagonal are all zeros ($l_{iis} = 0$). The entire adjacency matrix $\mathbf{L}$ is block diagonal where each block corresponds to a school s. From $\mathbf{L}$ is constructed a row-normalized adjacency matrix $\mathbf{G}$. That is, elements in each row of $\mathbf{G}$ are weighted such that $\sum_{j=1}^{M} g_{ijs} = 1$.

While links are symmetric and binary, they may be rank ordered. To get a handle on this, suppose, for instance, that there is a latent value v_{ijs} that each individual i assigns to his link to each other individual j. Assume further that this is symmetric, so $v_{ijs} = v_{jis}$. A link exists whenever $v_{ijs} > 0$. Define i's ranking of his link to j as follows: $R_{ijs} = 1\{v_{ijs} > 0\} \sum_{k \neq i} 1\{v_{iks} \geq v_{ijs}\}$. So, $R_{ijs} = 1$ for i's closest friend, $R_{ijs} = 2$ for his next closest, etc. $R_{ijs} = 0$ if i and j are not linked.

For each school s and $k \in \{1, ..., M\}$, where M is the maximum school size, I construct a censored adjacency matrix $\mathbf{L^{cens}_{s,k}}$. Each element of $\mathbf{L^{cens}_{s,k}}$ is generated according to the

[1] There is a limited number of papers, mostly theoretical, that allow for continuous links (see, e.g., Bloch and Dutta, 2009; Baumann, 2016; Griffith, 2017).

[2] To keep the discussion simple, I focus on symmetric networks here. The qualitative case is qualitatively similar.

following rule:

$$l_{ijs,k}^{cens} = 1\{R_{ijs} > 0\}1\{\min\{R_{ijs}, R_{jis}\} \leq k\} \tag{2.1}$$

That is, $l_{ijs,k}^{cens} = 1$ if *either* lists the other among her first k links. In this way, it is possible for an individual to have more than k observed links when the data is generated with a censoring rule of k. $l_{ijs,k}^{cens} = 0$ in one of two situations: either i and j are not friends or they are friends and their link is censored.

By combining all S of the matrices $\mathbf{L_{s,k}^{cens}}$ into a single block diagonal matrix, we construct $\mathbf{L_k^{cens}}$ for each $k \leq M$. With any censoring rule k, we only observe the censored $\mathbf{L_k^{cens}}$ rather than the true $\mathbf{L_k}$ whenever $k < M$. From this we construct a row-normalized censored adjacency matrix $\mathbf{H_k}$.

2.3.1.2 Outcomes Conditional on the Network

Outcomes are determined according to the linear-in-means process specified by Manski (1993). That is, outcomes are determined according to Equation 2.2.[3]

$$y_{is} = \beta_0 + \beta_1 \bar{y}_{is} + \beta_2 x_{is} + \beta_3 \bar{x}_{is} + \epsilon_{is} \tag{2.2}$$

In this equation, y_{is} is some outcome for individual i in school s and x_{is} is some characteristic for the same individual. Peer effects enter through $\bar{y}_{is}$, the mean of individual i's links' outcome, and $\bar{x}_{is}$, the mean of those same peers' exogenous characteristics. Rewritten in matrix form, this becomes

$$\mathbf{y} = \beta_0 \iota + \beta_1 \mathbf{G} \mathbf{y} + \beta_2 \mathbf{x} + \beta_3 \mathbf{G} \mathbf{x} + \epsilon \tag{2.3}$$

[3] Here, I assume that the coefficients are constants. I discuss issues that arise when coefficients are non-constant in Section 2.7 below.

where $\mathbf{y}$ is the vector of all outcomes y_{is}, $\mathbf{y}$ is a matrix of covariates x_{is}, and $\mathbf{G}$ is the row-normalized adjacency matrix defined in the previous subsection.

Assumption 2.1. $\mathbb{E}[\epsilon_{is}|\mathbf{x}, \mathbf{G}] = 0$

Assumption 2.1 provides the primary exogeneity assumption that is maintained throughout this paper. This assumes independence of unobservables ϵ_{is} from observables for i and all others. Further, and crucially, it assumes network endogeneity. That is, unobservables that play a part in forming networks $\mathbf{G}$ are not correlated with unobservables in the outcome equation ϵ. I acknowledge that network exogeneity is a crucial limitation. However, in the rest of this paper, I demonstrate that, even with this very strong assumption on the data generating process, estimates of parameters of the model in Equations (2.2) and (2.3) are still biased due to censoring.

2.3.2 Estimators and Uncensored Results

Here, I describe the two estimators that are used throughout this paper. These two estimators are motivated by criticisms of the linear-in-means model in Equation 2.2 that date back to Manski (1993), particularly the problem of reflection. I demonstrate that these estimators are consistent when there is no censoring of the data.

2.3.2.1 Reduced-Form OLS (Estimator 1)

Given reflection and the consequent failure to separately identify β_1, β_2, and β_3 in Equation 2.2, many researchers have simply estimated a reduced-form version by OLS. Assume for simplicity that y_{is}, x_{is}, and $\bar{x}_{is}^{true}$ have been demeaned (and thus $\beta_0 = 0$). Manipulation of Equation 2.3 yields

$$\mathbf{y} = (I - \beta_1 \mathbf{G})^{-1}(\beta_2 \mathbf{x} + \beta_3 \mathbf{G}\mathbf{x} + \epsilon) \tag{2.4}$$

Define the reduced-form OLS estimator $\hat{\alpha}$ as a regression of outcome y_{is} on regressors x_{is} and $\bar{x}_{is}$. That is, in matrix form,

$$\hat{\alpha} = \left([\mathbf{x}, \mathbf{G}\mathbf{x}]'[\mathbf{x}, \mathbf{G}\mathbf{x}]\right)^{-1} [\mathbf{x}, \mathbf{G}\mathbf{x}]'\mathbf{y} \tag{2.5}$$

Proposition 2.1. *Given Assumption 2.1,*

$$\text{plim}\left(\begin{bmatrix}\hat{\alpha}_1 \\ \hat{\alpha}_2\end{bmatrix}\right) = \begin{bmatrix}\beta_2 \\ \beta_3 + \beta_1\beta_2\end{bmatrix} + (\beta_3 + \beta_1\beta_2)\sum_{k=1}^{\infty}\beta_1^k \left(\mathbb{E}[[\mathbf{x}, \mathbf{G}\mathbf{x}]'[\mathbf{x}, \mathbf{G}\mathbf{x}]]\right)^{-1}\begin{bmatrix}\mathbb{E}[\mathbf{x}'\mathbf{G}^{k+1}\mathbf{x}] \\ \mathbb{E}[\mathbf{x}'\mathbf{G}'\mathbf{G}^{k+1}\mathbf{x}]\end{bmatrix}$$

Similar to others (see, e.g., Carrell, Sacerdote and West, 2013), I define the reduced form parameter $\alpha = \text{plim}\,\hat{\alpha}$, where $\text{plim}\,\hat{\alpha}$ is given in Proposition 2.1. In general, this reduced-form parameter α depends on two features of the data-generating process. First, it depends on the structural parameters $\beta = (\beta_1, \beta_2, \beta_3)$. Second, α depends on the relationships among exogenous x_{is}, peer group mean $\bar{x}_{is}$, and higher-order means $\bar{x}_{is}^k$. These relationships do not admit a simple closed form, as they are complicated expressions of the network-formation process, and how x_{is} plays a role in determining links. For example, if there is homophily in link formation conditional on x_{is}, the mean of individual's peers' characteristics $\bar{x}_{is}$ will be correlated with his own characteristics x_{is}. Therefore, $\mathbb{E}[x_{is}\bar{x}_{is}] > 0$ (in matrix notation, $\mathbb{E}[\mathbf{x}'\mathbf{G}\mathbf{x}] > 0$). In the simulations that follow, I approximate α numerically, given β and the parameters that lead to network formation.

The definition of α provides a generalization of the standard sufficiency result for the existence of peer effects. From the structural model, peer effects exist if *either* $\beta_1 \neq 0$ or $\beta_3 \neq 0$. From Proposition 2.1, we see that a sufficient condition for this is that $\alpha_2 \neq 0$. Note that this is a one-way implication, and there are combinations of parameters such that $\alpha_2 = 0$ yet peer effects still are present.

Two special cases bear mentioning at this point. First, when there is no endogneous peer effect ($\beta_1 = 0$), the reduced-form model is the true model, and $(\alpha_1, \alpha_2) = (\beta_2, \beta_3)$. In this case, it is easy to see that $\text{plim}(\hat{\alpha}_1, \hat{\alpha}_2) = (\beta_2, \beta_3)$.

Corollary 2.1. $\mathbf{G} = \mathbf{G}^2 \;\Rightarrow\; \text{plim}\left(\begin{bmatrix} \hat{\alpha}_1 \\ \hat{\alpha}_2 \end{bmatrix}\right) = \begin{bmatrix} \beta_2 \\ \frac{\beta_3+\beta_1\beta_2}{1=\beta_1} \end{bmatrix}$

Second, Corollary 2.1 gives a result that relates the expression from Proposition 2.1 to known results that map the parameters of the full model to the reduced form.(See, e.g., Carrell, Sacerdote and West, 2013). A well-known case of when $\mathbf{G} = \mathbf{G}^2$ is a classroom setting in which all members (including oneself) of the classroom are assumed to have equal weights.

2.3.2.2 Instrumental Variables Using Second-Order Peers (Estimator 2)

While reflection has been noted and dealt with since Manski (1993), the non-identification result is a knife-edge case (see, e.g., Blume et al., 2015). When $\mathbf{G} \neq \mathbf{G}^2$, second-order peers (friends of friends) may be used to surmount the reflection problem (Bramoullé, Djebbari and Fortin, 2009; DeGiorgi, Pelllizzari and Radaelli, 2010). Others have shown that variation in peer group size can also overcome reflection (Lee, 2007; Boucher et al., 2014). The former strategy amounts to, essentially, using the mean of friends' mean exogenous characteristics as an instrument for necessarily endogenous $\bar{y}_{is}$. In matrix terms, we employ $\mathbf{G}^2\mathbf{x}$ as an instrument for $\mathbf{Gy}$. The availability of this instrument (and higher-order instruments) necessarily depends on $\mathbf{G}^2 \neq \mathbf{G}$.

Define the following matrices:

1. $\mathbf{X} = [\mathbf{Gy}, \mathbf{x}, \mathbf{Gx}]$

2. $\mathbf{Z} = [\mathbf{x}, \mathbf{Gx}, \mathbf{G}^2\mathbf{x}]$

Proposition 2.2. *Given Assumption 2.1, if the matrix* $\mathbf{X}'\mathbf{ZZ}'\mathbf{X}$ *is invertible, then*

$$\text{plim}(\begin{bmatrix}\hat{\beta}_1\\ \hat{\beta}_2\\ \hat{\beta}_3\end{bmatrix}) = (\mathbf{X}'\mathbf{ZZ}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{ZZ}'\mathbf{y} = \begin{bmatrix}\beta_1\\ \beta_2\\ \beta_3\end{bmatrix}$$

The assumption that the matrix is invertible is not innocuous. Situations in which it fails, as pointed out by, e.g., Bramoullé, Djebbari and Fortin (2009), include when $\mathbf{G}^2 = \mathbf{G}$, or when certain combinations of parameters are zero. In the former case, the excluded instrument $\mathbf{G}^2\mathbf{x}$ is collinear with included regressor $\mathbf{Gx}$.[4] In the latter, $\mathbf{G}^2\mathbf{x}$ is uncorrelated with $\mathbf{Gy}$.

2.3.3 Censoring-Induced Bias

Having provided expressions for the probability limits for two common estimators in the absence of censoring, here I show how censoring can lead to deviations from these theoretical results. Recall from Subsection 2.3.1 above that censoring arises when $k < M$, where M is the maximum number of students in a school. When this is the case, censoring leads to observation of the censored row-normalized adjacency matrix $\mathbf{H_k}$ rather than the true row-normalized adjacency matrix $\mathbf{G}$. Clearly, data is uncensored whenever $\mathbf{H_k} = \mathbf{G}$.

2.3.3.1 Estimator 1

Here, I provide an analytic expression for the bias in Estimator 1 that is induced by censoring. For a given censoring rule k, define

$$\hat{\alpha}^{cens,k} = ([\mathbf{x}, \mathbf{H_k x}]'[\mathbf{x}, \mathbf{H_k x}])^{-1}[\mathbf{x}, \mathbf{H_k x}]'\mathbf{y} \tag{2.6}$$

[4] Indeed, this is exactly the case addressed by Manski (1993).

That is, $\hat{\alpha}^{cens,k}$ is the analogue of $\hat{\alpha}$ when the mean is calculated with censoring at a maximum of k links. Equation 2.7 gives the probability limit of this estimator when the number of schools/networks approaches infinity.

$$\begin{aligned}
\text{plim}(\hat{\alpha}^{cens,k}) &= \begin{bmatrix} \mathbb{E}[\mathbf{x}'\mathbf{x}] & \mathbb{E}[\mathbf{x}'\mathbf{H_k}\mathbf{x}] \\ \mathbb{E}[\mathbf{x}'\mathbf{H_k}\mathbf{x}] & \mathbb{E}[\mathbf{x}'\mathbf{H_k}'\mathbf{H_k}\mathbf{x}] \end{bmatrix}^{-1} \begin{bmatrix} \mathbb{E}[\mathbf{x}'\mathbf{x}] & \mathbb{E}[\mathbf{x}'\mathbf{H_k}\mathbf{x}] \\ \mathbb{E}[\mathbf{x}'\mathbf{G}\mathbf{x}] & \mathbb{E}[\mathbf{x}'\mathbf{H_k}'\mathbf{G}\mathbf{x}] \end{bmatrix} \alpha \\
&= \mathbf{B}_\alpha \alpha \qquad (2.7)
\end{aligned}$$

Clearly, this estimator does not converge to α whenever $\mathbf{B}_\alpha$ is not the identity matrix. From the expression in Equation 2.7, we see that the bias depends on the relationship between the variance of the censored measures and the covariance between the censored ($\mathbf{H_k x}$) and uncensored ($\mathbf{Gx}$) measures.

Two key features bear noting here. First, if we define measurement error as the difference $\mathbf{H_k x} - \mathbf{Gx}$, this formula is merely a special case of the general formula for measurement error that relates inconsistency of OLS in the presence of measurement errors (See Bound, Brown and Mathiowetz, 2001). Second, as k grows, $\mathbf{H_k} \to \mathbf{G}$ and thus $\mathbf{B}_\alpha \to \mathbf{I}$. Therefore, as expected, $\lim_{k \to M} \text{plim}(\hat{\alpha}^{cens,k}) = \alpha$. So, as we allow individuals to name more links, we get less censoring, until eventually there is no censoring-induced bias in the limit.

2.3.3.2 Estimator 2

Here I show that measurement error due to censoring generates bias in Estimator 2. Define the following matrices.

1. $\mathbf{X_k^{cens}} = [\mathbf{H_k y}, \mathbf{x}, \mathbf{H_k x}]$

2. $\mathbf{Z_k^{cens}} = [\mathbf{x}, \mathbf{H_k x}, \mathbf{H_k}^2 \mathbf{x}]$

Next, define the estimator

$$\hat{\beta}^{cens,k} = (\mathbf{X_k^{cens\prime}} \mathbf{Z_k^{cens}} \mathbf{Z_k^{cens\prime}} \mathbf{X_k^{cens}})^{-1} \mathbf{X_k^{cens\prime}} \mathbf{Z_k^{cens}} \mathbf{Z_k^{cens\prime}} \mathbf{y} \tag{2.8}$$

That is, $\hat{\beta}^{cens,k}$ is the analogue to $\hat{\beta}^{uncens}$ that is estimated using censored versions of all network variables (constructed from $\mathbf{H_k}$ instead of $\mathbf{G}$). Equation 2.9 gives the convergence result for $\hat{\beta}^{cens,k}$.

$$\begin{aligned}
\text{plim}(\hat{\beta}^{cens,k}) = & \begin{bmatrix} \mathbb{E}[(\mathbf{H_k}^2\mathbf{x})'\mathbf{H_k}\mathbf{y}] & \mathbb{E}[(\mathbf{H_k}^2\mathbf{x})'\mathbf{x} & \mathbb{E}[(\mathbf{H_k}^2\mathbf{x})'\mathbf{H_k}\mathbf{x}] \\ \mathbb{E}[\mathbf{x}'\mathbf{H_k}\mathbf{y}] & \mathbb{E}[\mathbf{x}'\mathbf{x}] & \mathbb{E}[\mathbf{x}'\mathbf{H_k}\mathbf{x}] \\ \mathbb{E}[(\mathbf{H_k}\mathbf{x})'\mathbf{H_k}\mathbf{y}] & \mathbb{E}[(\mathbf{H_k}\mathbf{x})'\mathbf{x}] & \mathbb{E}[(\mathbf{H_k}\mathbf{x})'\mathbf{H_k}\mathbf{x}] \end{bmatrix}^{-1} \\
& \quad \begin{bmatrix} \mathbb{E}[(\mathbf{H_k}^2\mathbf{x})'\mathbf{G}\mathbf{y}] & \mathbb{E}[(\mathbf{H_k}^2\mathbf{x})'\mathbf{x} & \mathbb{E}[(\mathbf{H_k}^2\mathbf{x})'\mathbf{G}\mathbf{x}] \\ \mathbb{E}[\mathbf{x}'\mathbf{G}\mathbf{y}] & \mathbb{E}[\mathbf{x}'\mathbf{x}] & \mathbb{E}[\mathbf{x}'\mathbf{G}\mathbf{x}] \\ \mathbb{E}[(\mathbf{H_k}\mathbf{x})'\mathbf{G}\mathbf{y}] & \mathbb{E}[(\mathbf{H_k}\mathbf{x})'\mathbf{x}] & \mathbb{E}[(\mathbf{H_k}\mathbf{x})'\mathbf{G}\mathbf{x}] \end{bmatrix} \beta \\
= & \; \mathbf{B}_\beta \beta
\end{aligned} \tag{2.9}$$

As in the case of $\hat{\alpha}^{k,cens}$, the bias depends on the relationship between censored and uncensored versions of variables. However, unlike $\hat{\alpha}^{k,cens}$, this includes endogenous variables as well as exogenous ones (including assumed exogenous networks). Further, it is easy to see that in the limiting case where $\mathbf{G} = \mathbf{H_k}$ (no censoring), $\mathbf{B}_\beta$ is the identity matrix, and $\hat{\beta}^{cens,k} = \hat{\beta}$ is a consistent estimator of β.

While IV estimators have been heavily used to fix measurement error, in this case the standard result does not hold. This is due to the fact that censoring induces measurement error in *both* both $\mathbf{Gy}$ (the endogenous regressor) and $\mathbf{G}^2\mathbf{x}$ (the excluded instrument). Instead of observing these "true" variables, we observe censored versions of them in the form of $\mathbf{H_k y}$ and $\mathbf{H_k}^2\mathbf{x}$. As pointed out by Chandrasekhar and Lewis (2011), the measurement error in these two variables is correlated.

Further, censoring leads to failure of the exclusion restriction required for consistency of $\hat{\beta}$. This is due to the highly parametric assumption that motivates this estimator derived by Bramoullé, Djebbari and Fortin (2009) and DeGiorgi, Pelllizzari and Radaelli (2010). The exclusion restriction requires that second-order friends ("friends of friends") influence outcomes only through first-order friends and not directly. However, censoring of friendship networks leads to some first-order friends being incorrectly identified as second-order friends. This implies that, even if data is generated according to the highly restrictive linear-in-means model, the exclusion restriction required for consistency is not met when friendship networks are censored.

2.4 Simulations of Bias

To demonstrate inconsistency due to censoring, here I perform a series of Monte Carlo experiments. These experiments consist of two steps. First, data is generated according to a two-step process. Second, I estimate $\hat{\alpha}^{cens,k}$ and $\hat{\beta}^{cens,k}$ on the data for all values of $k \leq M$. From these simulations, we see that whenever k is small, estimates of $\hat{\alpha}^{cens,k}$ and $\hat{\beta}^{cens,k}$ are very far from the uncensored estimates, demonstrating that these estimates are sensitive to the number of links observed.

2.4.1 DGP

Here I describe the data generation process that I use for simulations. Final outcomes are generated according to a two-step process. First, a network is formed according to a simple dyadic network-formation process. Second, conditional on the realized network, outcomes are determined according to Equation (2.2).

2.4.1.1 A Simple Model of Exogenous Networks

As the purpose of this paper is not to test different models of network formation, in running simulations, I keep the network formation process simple. To allow for some homophily, I specify network formation as follows. A link between individuals i and j, denoted l_{ijs} exists according to the following rule:

$$l_{ijs} = \mathbf{1}\{\gamma_0 + |X_{is} - X_{js}|\gamma_1 + u_{ijs} > 0\} = \mathbf{1}\{U_{ijs} > 0\} \tag{2.10}$$

The distance function $|X_{is} - X_{js}|$ allows link probability to depend on how close individuals i and j are in characteristics space.

Assumption 2.2. *The following assumptions are made on the link-formation process*

1. $\gamma_1 \leq 0$ *(homophily)*
2. $u_{ijs} = u_{jis}$ *(links are symmetric)*
3. $i \neq k$ *or* $j \neq l \;\Rightarrow\; u_{ijs} \perp\!\!\!\perp u_{kls}$ *(unobservables are independent across links)*

For simplicity, I impose the restrictions in Assumption 2.2. Under these assumptions, a special case of this model, corresponding to the case when $\gamma_1 = 0$, is the Erdos-Renyi random graphs model, in which each link l_{ijs} is a draw from a Bernoulli(p) distribution where $p = \Pr(u_{ijs} > -\gamma_0)$. The more general model that allows p to vary based on observables is a special case of the stochastic block models that have received substantial attention in the literature.

From this data-generating process, each individual i can rank order her links by utility. So, she lists her first friend k such that $U_{iks} = \max_{l \neq i} U_{ils}$ and others in order until $U_{ils} < 0$. For a given number $k \leq M$, we construct a censored adjacency matrix $\mathbf{L}^{cens,k}$. In turn, for each k, I then construct a row-normalized $\mathbf{H_k}$.

In all simulations, the following assumptions are made

1. $X_{is} \overset{\text{i.i.d.}}{\sim} N(0, 1)$
2. $u_{ijs} = u_{jis} \overset{\text{i.i.d.}}{\sim} N(0, 1)$

3. 200 networks of $M = 25$ are observed ($N = 5000$)

I allow the following to vary:

1. $\gamma_1 \in \{0, -1, -2\}$

2. γ_0 such that $\Pr(l_{ijs} = 1) \in \{0.20, 0.35, 0.50\}$

Different γ_1 varies the importance of $|X_{is} - X_{js}|$ in the link formation process and thus varies the strength of homophily. Different γ_0 varies the network density. Accordingly, for each repetition and each $k \leq M$, I simulate nine sets of estimates $\hat{\alpha}^{cens,k}$ and $\hat{\alpha}^{cens,k}$, corresponding to each combination of values for γ_1 and γ_0.

2.4.1.2 Outcomes Conditional on the Network

Conditional on the true network $\mathbf{G} = \mathbf{H_M}$, outcomes are determined according to Equation 2.2. The matrices $\mathbf{G}$ and $\mathbf{x}$ are determined in simulating the network-formation process. In all simulations, the following is set

1. $\epsilon_{is} \overset{\text{i.i.d.}}{\sim} N(0, 1)$

2. $\beta' = (\beta_0, \beta_1, \beta_2, \beta_3)' = (\ 0,\ 0.6,\ 1,\ 0.5\ )$

In practice, due to reflection, I construct outcomes $\mathbf{y}$ by Equation (2.11), which is a rearranged version of Equation (2.3).

$$\mathbf{y} = (I - \beta_1 \mathbf{G})^{-1}(\beta_2 \mathbf{x} + \beta_3 \mathbf{G}\mathbf{x} + \epsilon) \tag{2.11}$$

2.4.2 Simulated Estimates

Here, I present estimates of the extent of censoring-induced bias for both estimators. From this, we see that, in general, censoring leads to biased estimates for both estimators, with predictable exceptions. These results lend support to the analytic results derived in the prior section, further suggesting that censoring may lead to crucially biased estimates of the parameters of peer effects models.

First, I look at bias in $\hat{\alpha}_1$ and $\hat{\alpha}_2$. As an initial matter, recall that the measurement error is in measuring $\mathbf{Gx}$, leading to biased estimates of $\hat{\alpha}_2$. The regressor for which α_1 is a coefficient, $\mathbf{x}$ is not measured with error. However, to the extent that $\mathbf{x}$ is correlated with $\mathbf{Gx}$, we also see bias in estimates of $\hat{\alpha}_1$.

Figure 2.1 conforms to expectations. Subfigure (a) presents the mean of 1000 simulations of 200 schools of 25 students each (N = 5000). In Figure 2.1, the unconditonal probability of a link in any given school is 0.35. Note first that, when $\gamma_1 = 0$, indicating no homophily, there is no bias due to censoring. This conforms to known results, since in this case the regressor x_{is} is uncorrelated with the measured-with-error regressor $\bar{x}_{is}$. When $\gamma_1 < 0$, we see that, for low quantities of K (the maximum amount of friends permitted), we see high amounts of bias, and these estimates are on average biased upward.

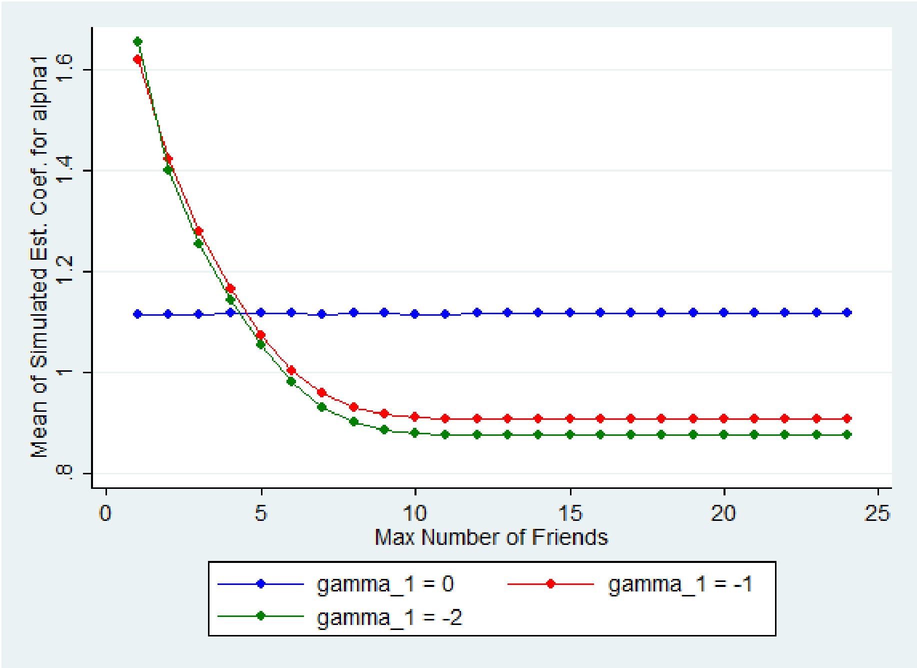
Subfigure (b) of Figure 2.1 shows that, at low K, there is substantial bias in mean estimated $\hat{\alpha}_2$. This bias is toward zero, as is the case with traditional attenuation bias. Note that, similar to the formula from Bound, Brown and Mathiowetz (2001), the bias in $\hat{\alpha}_2$ and $\hat{\alpha}_1$ are in opposite directions whenever $\gamma_1 < 0$. That is, when the measured-with-error variable and the correlated variable are positive correlated, their biases are in opposite directions. This same pattern shows up in the actual empirical estimates in the sections that follow.

Moving to Estimator 2 ($\hat{\beta}$), censoring similarly causes bias in simulations. These results are shown in Figure 2.2. Note that, for small values of K, all three estimates are biased except in the case of $\hat{\beta}_2$ when $\gamma_1 = 0$.

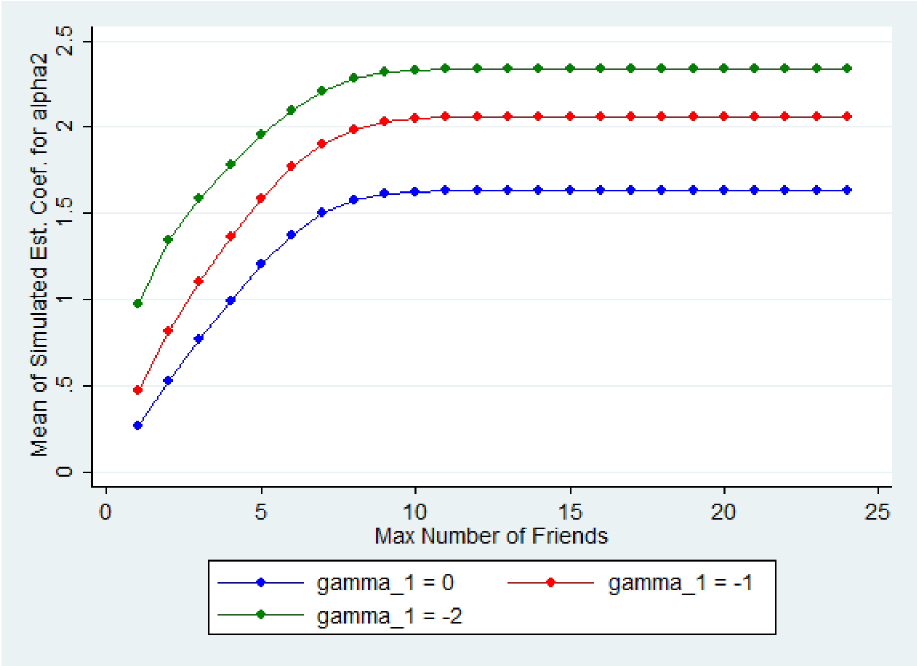
In sum, these results show that estimates of $\hat{\alpha}$ and $\hat{\beta}$ may be biased whenever there is substantial censoring in the observed friendship data. The mean of all simulated estimates converges to the true value of the parameter as we allow respondents to name more and more links, as shown in Figures 2.1 and 2.2.

Figure 2.1: Mean of Simulated $\hat{\alpha}$

(a) Mean of Simulated $\hat{\alpha}_1$



(b) Mean of Simulated $\hat{\alpha}_2$



Additionally, the amount of bias depends upon the amount of links that are censored. For a given K, the quantity of links censored is inceasing in network density. Accordingly, in AppendixE, Tables E.1 through E.5, I present additional results allowing the unconditional probability of a link existing to be, alternatively, 0.20, 0.35, or 0.50. From this, we see that, for a given censoring rule K, there is more bias in the estimates when the network is denser (and thus more links are censored).

2.5 Bias in Real Datasets

2.5.1 AddHealth

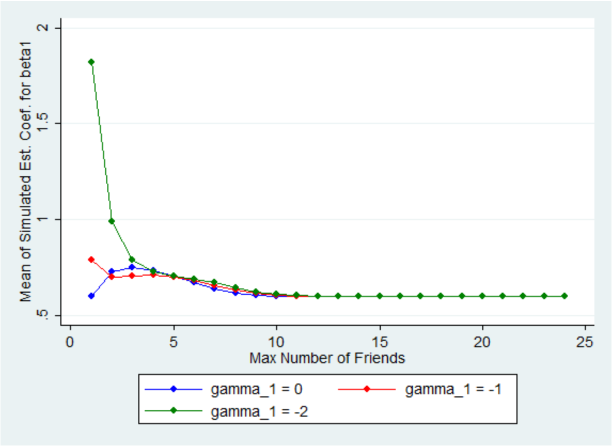
2.5.1.1 Empirical Strategy

The National Longitudinal Study of Adolescent to Adult Health (AddHealth) is a long-term panel study collected by researchers at the University of North Carolina. Throughout the course of the study, they collect detailed data on demographics and a number of academic and behavioral outcomes. Importantly, during the course of Wave 1 of the study, they also collect detailed social network data. Crucially, as discussed in Section 2.2 above, they only allow individuals to name up to five male and five female friends. As shown in Table 2.1, there is substantial scope for censoring in the data, as a large percentage of individuals name the maximum number of male and/or female friends.

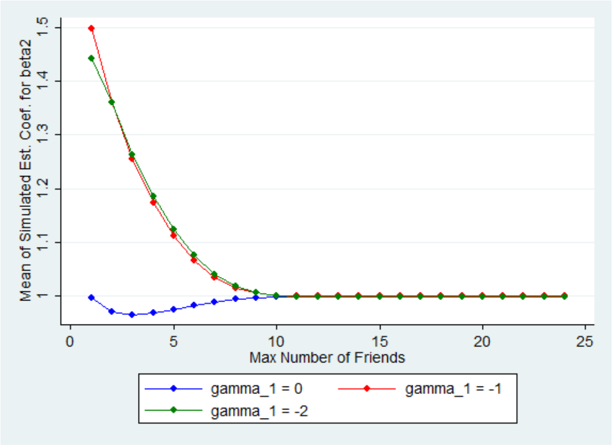
Given the censoring induced by the survey design, in AddHealth we never actually observe the "true" network. Thus, I cannot estimate $\hat{\alpha}$ or $\hat{\beta}$ with the full network data. Rather, I can estimate each wtih censoring at any number up to the maximum amount of friends allowed to be listed. That is, I estimate $\hat{\alpha}$ or $\hat{\beta}$ for $K \in \{1, 2, 3, 4, 5\}$. These results compare to the simulations in Figures 2.1 and 2.2, except that we only observe estimates for $K \leq 5$.

Figure 2.2: Mean of Simulated $\hat{\beta}$

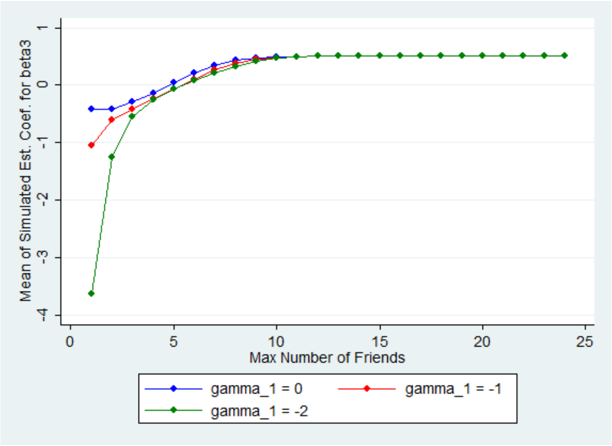
(a) Mean of Simulated $\hat{\beta}_1$



(b) Mean of Simulated $\hat{\beta}_2$



(c) Mean of Simulated $\hat{\beta}_3$



2.5.1.2 Data Description

The AddHealth data has been employed to study the associations between peers and a wide variety of academic and behavioral outcomes. For the purpose of this empirical exercise, I look at five academic and three behavioral outcomes as follows. These are summarized in Table 2.2, Panel A. For each outcome, I estimate $\hat{\alpha}$ and $\hat{\beta}$ for $K = 1, ..., 5$. Panel B summarizes the right-hand side variables that I use. Since there are nine such variables, for each outcome and every censoring rule K, $\hat{\alpha}_1$ and $\hat{\alpha}_2$ are 9-dimensional vectors.

Table 2.2: AddHealthVariables

	Min	Max
Panel A: Outcomes		
Grade Point Average in All Subjects	0	4
Grade Point Average in English	0	4
Grade Point Average in Math	0	4
Grade Point Average in History	0	4
Grade Point Average in Science	0	4
Indicator for Has Drank Alcohol	0	1
Indicator for Got Drunk	0	1
Indicator for Smoked	0	1
Panel B: Independent Variables		
Age	10	19
Grade	6	12
Hispanic	0	1
Black	0	1
Asian	0	1
Other Race	0	1
Born in the USA	0	1
Lives with Mother	0	1
Lives with Father	0	1

2.5.1.3 Estimates

Estimates of $\hat{\alpha}$ for the first outcome, GPA in All Subjects, are presented in Table 2.3. Each column corresponds to a different censoring rule: Column 1 uses only those indicated as first friends, Column 2 uses the first two, etc. Note that, as with the simulations in Section 2.4, the estimates on the peer mean variables ($\overline{\text{Age}}$, etc.) trend away from zero.

These estimates correspond to $\hat{\alpha}_2$.Correspondingly, the estimates on the raw variables (Age, etc.) trend toward zero as we move from 1 to 5 observed friends.

To conserve space, I present the coefficients in graphical form. Figure 2.3 plots the estimated coefficients for all outcomes in Table 2.2, plotted against censoring of 1 to 5 friends. This provides strong evidence that the pattern for GPA in All Subjects, shown in Table 2.3, holds for many covariates and outcomes. On the left are plotted coefficients for raw variables (Age, Grade, etc.), while the right column plots coefficients for corresponding peer mean variables ($\overline{\text{Age}}$, $\overline{\text{Grade}}$, etc.). While not universal, there is a strong trend away from zero for the latter, while the former tend to trend toward zero. These results are in strong agreement with the simulation results presented above.

Next, I present estimates of $\hat{\beta}$ for the first outcome, GPA in All Subjects, in Table 2.4. First, I note that estimated $\hat{\beta}_1$ is trending strongly upward as we observe more links. Figure 2.4 provides evidence that this trend holds for most of the eight outcomes.

Trends can be seen in the paths of coefficients $\hat{\beta}_2$ and $\hat{\beta}_3$ in Appendix E Figure E.1. While less clear than the trend for $\hat{\alpha}_1$, the coefficients on the demographic variables (Age, Grade, etc.) in the left side of Figure E.1 tend to be toward zero. Similarly, to the extent trends are apparent, the trends on the mean demographic variables on the right side of Figure E.1 tend to be away from zero as K increases.

In sum, this empirical exercise from AddHealth conforms to the predictions of the simulations above. While we do not observe the "true" values of the parameters, at low levels of K, there is substantial censoring of friendship netwroks. Due to this, we can see significant trends in the path of $\hat{\alpha}$ and $\hat{\beta}$ as K increases from 1 to 5.

2.5.2 Nepal Menstrual Cups Data

A worry with Addhealth is that the indepenent variables in the above regressions are likely not exogenous. Accordingly, caution should be used in interpreting any of the coefficients above as causal. In partial answer to that, I here present results from an experimental

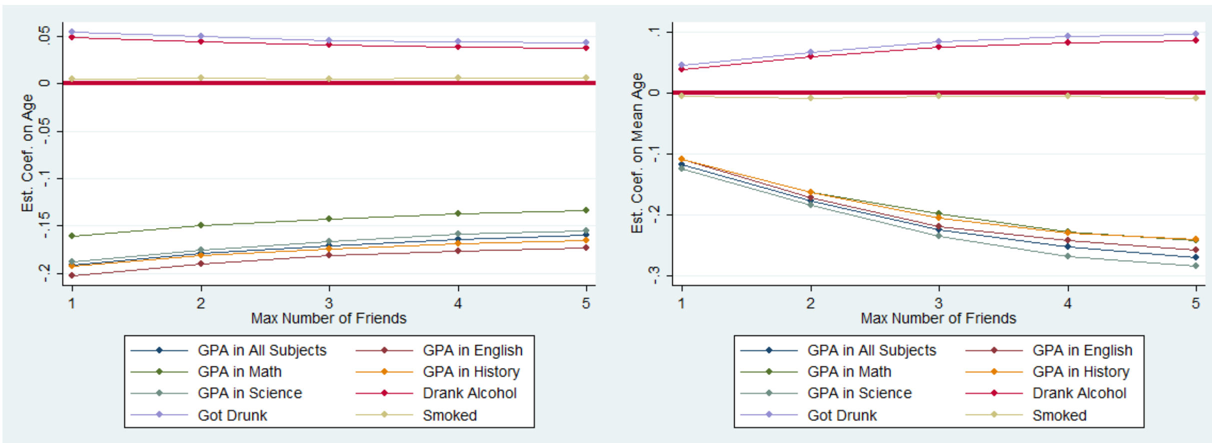
Table 2.3: AddHealth Regression Results for Estimator 1 (GPA in All Subjects)

	(1)	(2)	(3)	(4)	(5)
Age	-0.191***	-0.179***	-0.171***	-0.165***	-0.160***
	(0.013)	(0.013)	(0.012)	(0.012)	(0.012)
Grade	0.178***	0.166***	0.156***	0.156***	0.152***
	(0.016)	(0.017)	(0.018)	(0.018)	(0.019)
Hispanic	-0.182***	-0.154***	-0.137***	-0.132***	-0.127***
	(0.024)	(0.020)	(0.017)	(0.017)	(0.016)
Black	-0.166***	-0.125***	-0.105***	-0.095***	-0.084***
	(0.023)	(0.023)	(0.023)	(0.024)	(0.024)
Asian	0.174***	0.162***	0.147***	0.147***	0.142***
	(0.028)	(0.026)	(0.024)	(0.023)	(0.023)
Other Race	-0.063***	-0.055***	-0.053***	-0.052***	-0.048***
	(0.015)	(0.015)	(0.015)	(0.014)	(0.014)
Born in the USA	-0.078**	-0.093***	-0.096***	-0.098***	-0.097***
	(0.032)	(0.028)	(0.026)	(0.025)	(0.025)
$\overline{\text{Age}}$	-0.117***	-0.176***	-0.225***	-0.253***	-0.269***
	(0.013)	(0.020)	(0.024)	(0.027)	(0.029)
$\overline{\text{Grade}}$	0.097***	0.155***	0.203***	0.223***	0.238***
	(0.015)	(0.022)	(0.025)	(0.028)	(0.029)
$\overline{\text{Hispanic}}$	-0.138***	-0.167***	-0.190***	-0.196***	-0.199***
	(0.031)	(0.039)	(0.047)	(0.051)	(0.055)
$\overline{\text{Black}}$	-0.117***	-0.147***	-0.159***	-0.159***	-0.167***
	(0.032)	(0.040)	(0.044)	(0.046)	(0.047)
$\overline{\text{Asian}}$	0.093***	0.102**	0.118**	0.113*	0.117*
	(0.029)	(0.043)	(0.055)	(0.061)	(0.065)
$\overline{\text{Other Race}}$	-0.070***	-0.104***	-0.125***	-0.132***	-0.153***
	(0.023)	(0.032)	(0.037)	(0.041)	(0.045)
$\overline{\text{Born in the USA}}$	0.012	0.027	0.007	-0.004	-0.020
	(0.034)	(0.047)	(0.055)	(0.059)	(0.062)
Constant	4.566***	4.742***	4.928***	4.994***	5.022***
	(0.184)	(0.222)	(0.244)	(0.253)	(0.262)
Observations	32,315	32,315	32,315	32,315	32,315
R-squared	0.122	0.129	0.134	0.137	0.139

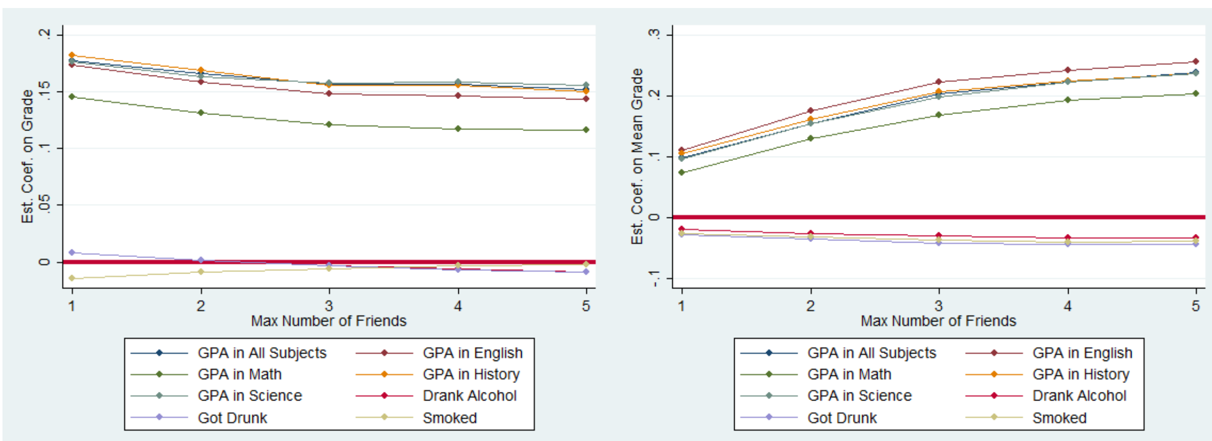
Notes: Standard errors in parentheses, clustered by school (138 schools). *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Coefficients for Lives with Mother, Lives with Father, $\overline{\text{Lives with Mother}}$, and $\overline{\text{Lives wtih Father}}$ not shown. Sample restricted to observations with non-missing data for all specifications.

Figure 2.3: Add Health Estimated Coefficients (Estimator 1)

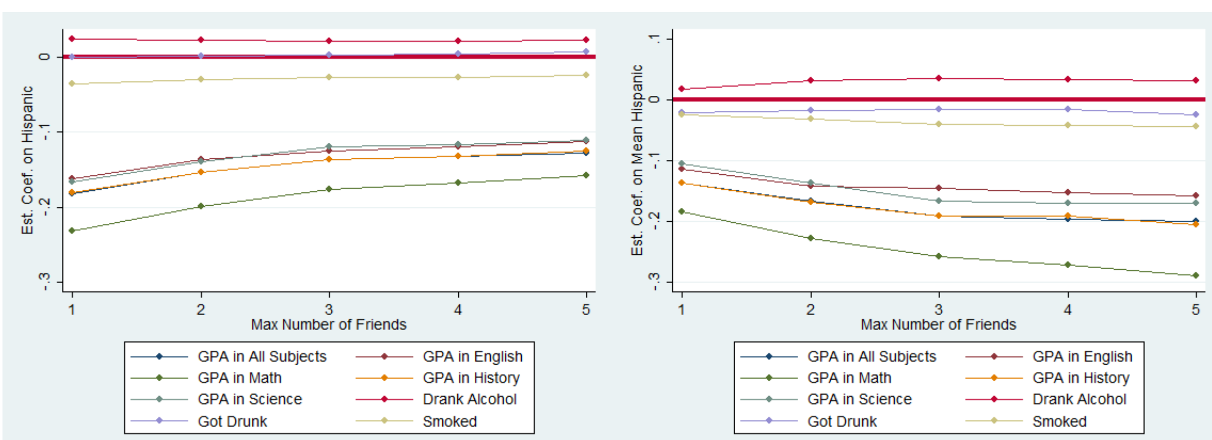
(a) Coefficients for Age and $\overline{\text{Age}}$



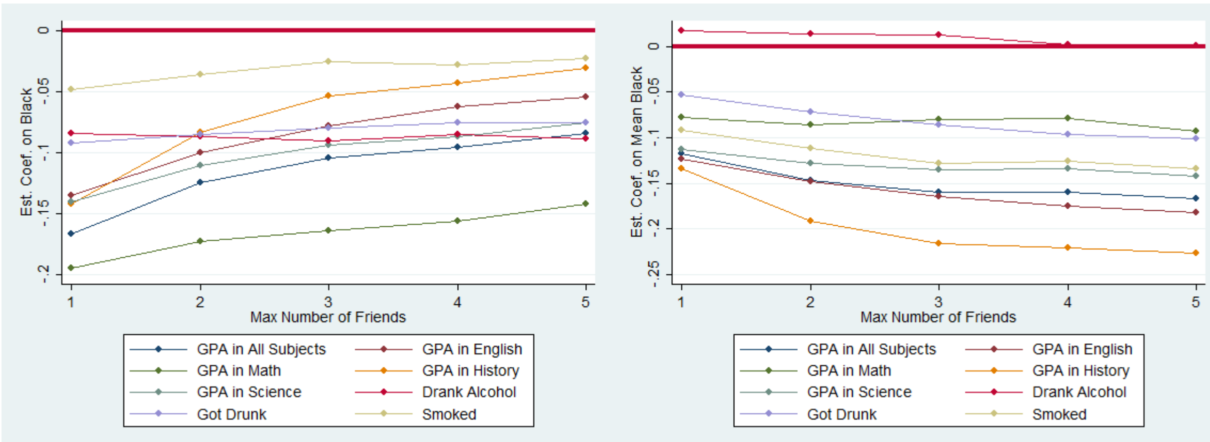
(b) Coefficients for Grade and $\overline{\text{Grade}}$



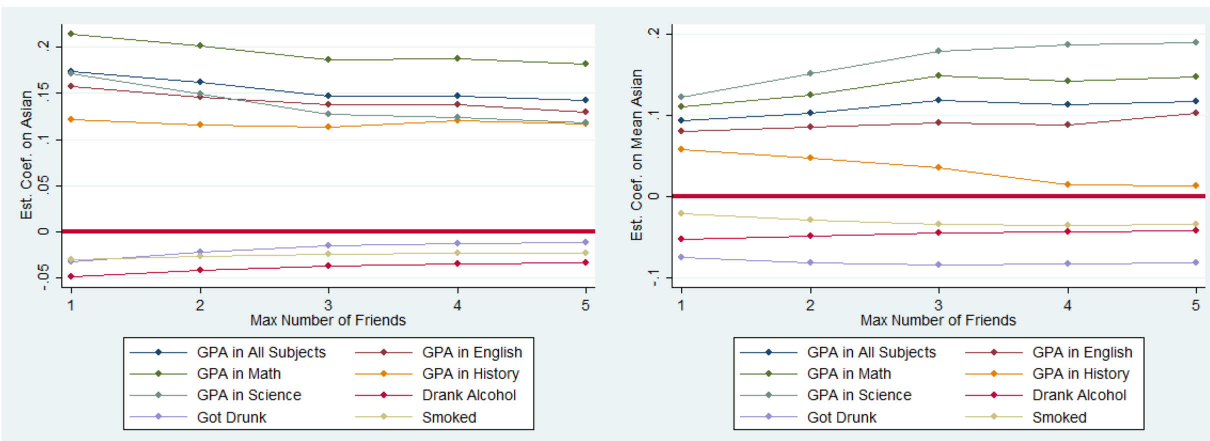
(c) Coefficients for Hispanic and $\overline{\text{Hispanic}}$



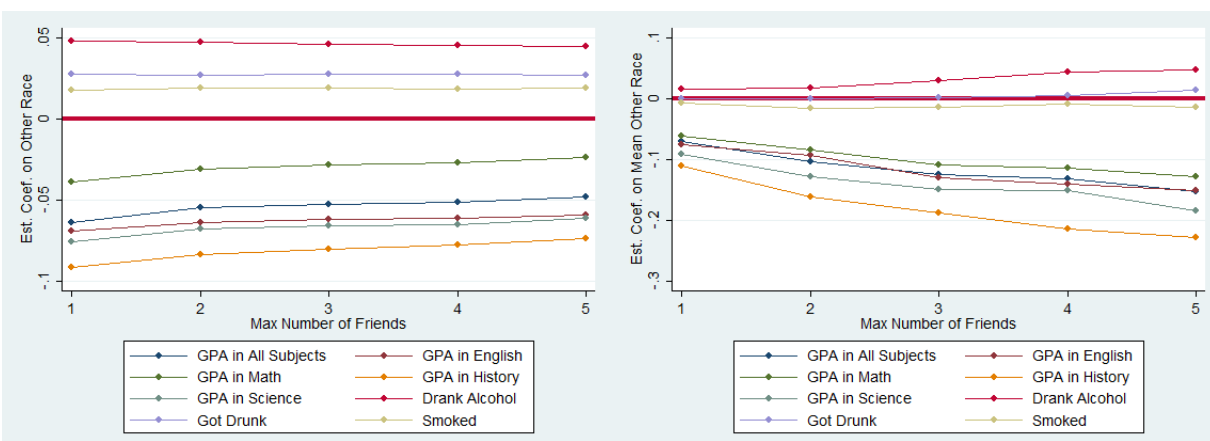
(d) Coefficients for Black and $\overline{\text{Black}}$



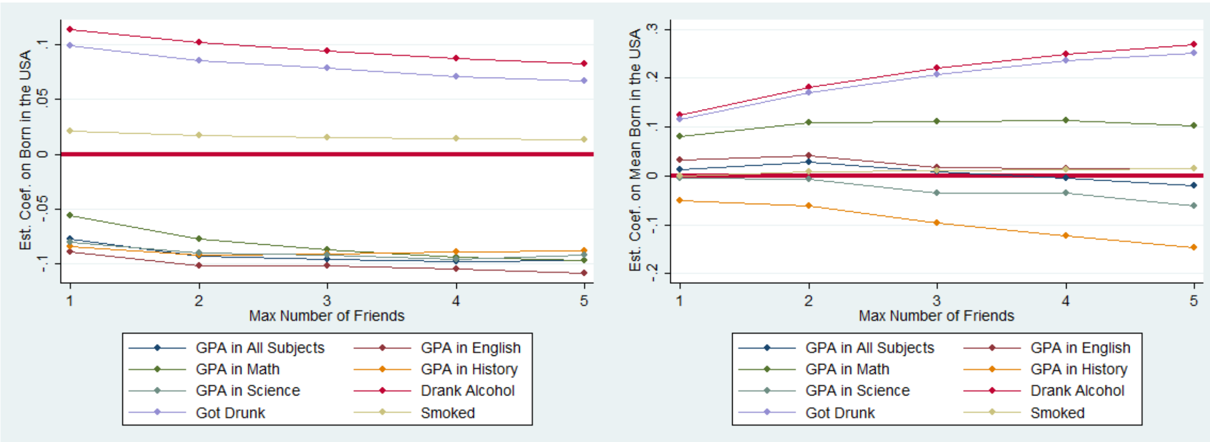
(e) Coefficients for Asian and $\overline{\text{Asian}}$



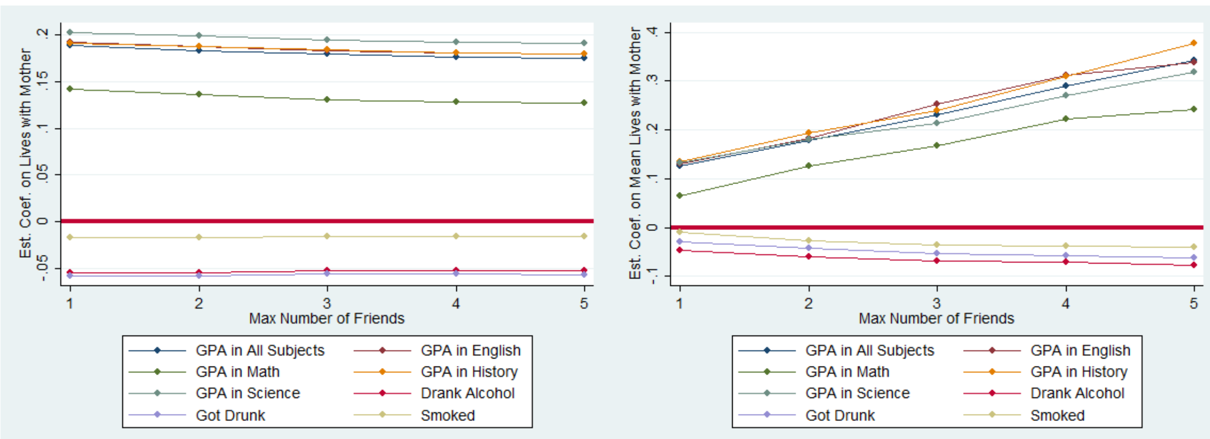
(f) Coefficients for Other Race and $\overline{\text{Other Race}}$



(g) Coefficients for Born in the USA and $\overline{\text{Born in the USA}}$



(h) Coefficients for Lives with Mother and $\overline{\text{Lives wtih Mother}}$



(i) Coefficients for Lives with Father and $\overline{\text{Lives with Father}}$

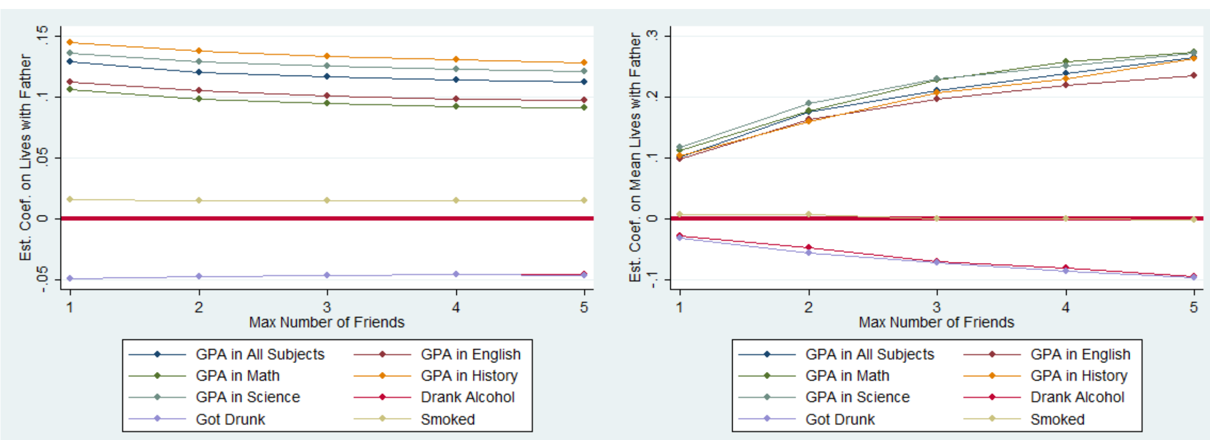
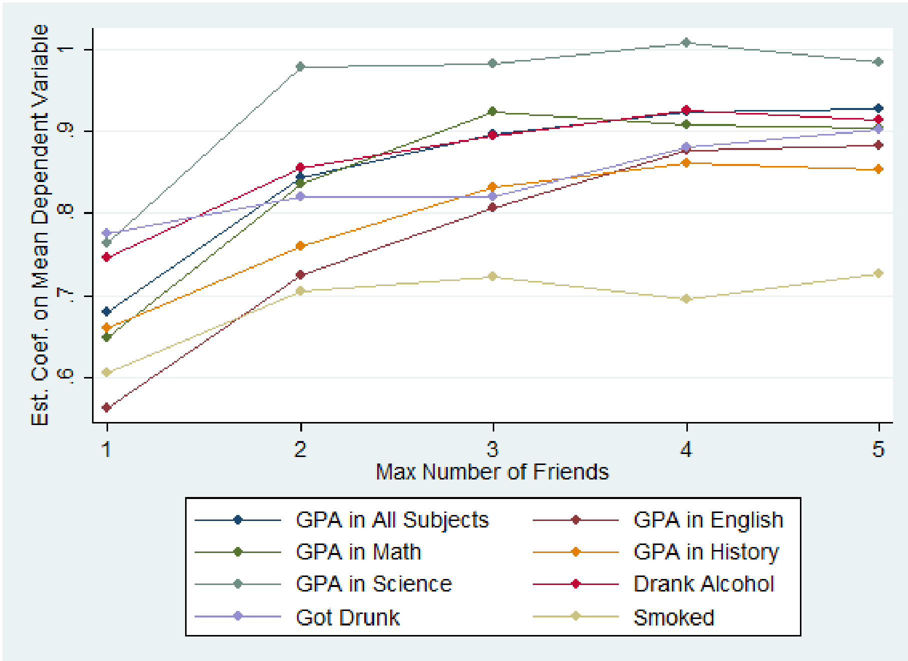


Table 2.4: AddHealth Regression Results for Estimator 2 (GPA in All Subjects)

	(1)	(2)	(3)	(4)	(5)
$\overline{\text{GPA in All Subjects}}$	0.681***	0.845***	0.896***	0.924***	0.928***
	(0.075)	(0.065)	(0.061)	(0.054)	(0.052)
Age	-0.138***	-0.119***	-0.117***	-0.115***	-0.113***
	(0.012)	(0.012)	(0.012)	(0.011)	(0.011)
Grade	0.119***	0.099***	0.092***	0.090***	0.087***
	(0.016)	(0.017)	(0.017)	(0.017)	(0.018)
Hispanic	-0.112***	-0.089***	-0.082***	-0.078***	-0.078***
	(0.018)	(0.017)	(0.017)	(0.017)	(0.017)
Black	-0.118***	-0.091***	-0.083***	-0.076***	-0.064**
	(0.027)	(0.028)	(0.027)	(0.027)	(0.027)
Asian	0.108***	0.105***	0.104***	0.111***	0.109***
	(0.021)	(0.021)	(0.023)	(0.023)	(0.023)
Other Race	-0.043**	-0.041**	-0.038**	-0.036**	-0.034**
	(0.017)	(0.016)	(0.016)	(0.015)	(0.015)
Born in the USA	-0.078***	-0.081***	-0.084***	-0.083***	-0.086***
	(0.025)	(0.024)	(0.023)	(0.023)	(0.023)
$\overline{\text{Age}}$	0.019	0.049**	0.046*	0.045*	0.034
	(0.023)	(0.025)	(0.025)	(0.026)	(0.026)
$\overline{\text{Grade}}$	-0.012	-0.035	-0.028	-0.026	-0.014
	(0.022)	(0.023)	(0.025)	(0.026)	(0.026)
$\overline{\text{Hispanic}}$	0.011	0.022	0.024	0.034	0.037
	(0.027)	(0.029)	(0.028)	(0.028)	(0.031)
$\overline{\text{Black}}$	0.020	0.048	0.056	0.066*	0.060*
	(0.034)	(0.037)	(0.036)	(0.036)	(0.035)
$\overline{\text{Asian}}$	-0.027	-0.078**	-0.101**	-0.109***	-0.102***
	(0.030)	(0.036)	(0.040)	(0.038)	(0.037)
$\overline{\text{Other Race}}$	-0.017	0.003	0.002	0.002	0.001
	(0.024)	(0.027)	(0.031)	(0.033)	(0.033)
$\overline{\text{Born in the USA}}$	0.045**	0.041	0.028	0.034	0.038
	(0.022)	(0.028)	(0.030)	(0.033)	(0.034)
Constant	1.555***	0.806**	0.652**	0.532*	0.537**
	(0.346)	(0.315)	(0.301)	(0.272)	(0.263)
Observations	27,740	27,740	27,740	27,740	27,740
R-squared	0.178	0.216	0.257	0.279	0.295

Notes: Standard errors in parentheses, clustered by school (138 schools). *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Coefficients for Lives with Mother, Lives with Father, $\overline{\text{Lives with Mother}}$, and $\overline{\text{Lives wtih Father}}$ not shown. Sample restricted to observations with non-missing data for all specifications. Excluded instruments for endogenous $\overline{\text{GPA in All Subjects}}$ are means of second-order links (Bramoullé, Djebbari and Fortin, 2009; DeGiorgi, Pelllizzari and Radaelli, 2010).

Figure 2.4: Add Health Estimated Coefficients on Mean Dependent Variable (Estimator 2)



intervention where the right-hand variables were assigned randomly.

2.5.2.1 Project Description

Many barriers exist to girls' schooling in the developing world. One often cited by policymakers is menstruation and the need for sanitary products. Accordingly, in an effort to test wehther relieving this barrier would help girls attend school, Oster and Thornton (2011) conducted a randomized trial of an intervention in rural Nepal.

The intervention consisted of randomly assigning girls to treatment and control groups. Girls in the treatment group received free sanitary products (a menstrual cup) as well as instruction on use. Importantly for purposes here, the study team collected detailed data on social connections among girls in the study schools. They asked each girl to name her friends, up to a maximum of three. From this they were able to study how this new technology was adopted through the social network (Oster and Thornton, 2012).

2.5.2.2 Empirical Strategy

Due to data limitations, I do not estimate $\hat{\beta}$ here. Rather, I focus on $\hat{\alpha}$. As discussed earlier, a large fraction of the girls in the survey named the maximum number of links allowed (three). Accordingly, as in AddHealth, we do not directly observe the entire social network and can only calculate censored versions of $\hat{\alpha}$ and look for trends.

I focus on two outcomes that are central to the analysis in Oster and Thornton (2012). The first is whether a girl tried to use the menstrual cup. The second is whether she successfully used the product. See the prior paper for a fuller description of these variables and their construction.

Further, since only girls in the treatment group had access to a menstrual cup, only girls in the treatment group could try to use and successfully use the product. Therefore, I estimate Equation (2.12) within the treatment group:

$$y_{ist} = \alpha_0 + \alpha_2 \overline{Treat}_{is} + \epsilon_{is} \tag{2.12}$$

where y_{ist} is an outcome (Tried or Used) for girl i in school s at time t. $\overline{Treat}_{is}$ is the fraction of a girl's friends who are also assigned to the treatment group. Note that, since treatment is randomly assigned within a school, we can be confident that $\overline{Treat}_{is}$ is independent in expectation of unobserved ϵ_{is}. I estimate $\hat{\alpha}_2$ in Equation (2.12) calculating $\overline{Treat}_{is}$ censored at 1, 2, or 3.

2.5.2.3 Results

Here, I present estimates of the parameters of Equation (2.12) for different censoring rules. From this, we see that estimates of $\hat{\alpha}_2$ trend strongly away from zero as we move from 1 to 3 observed friends. This provides further support to the analytic and simulated results.

Oster and Thornton (2012) collected monthly outcome data for all individuals at 13 points in time. Accordingly, I present estimates of censored $\hat{\alpha}_2$ in two different ways. First, I pool all months together in Table 2.5. Odd-numbered columns are specifications using only $\overline{Treat_{is}}$ and school fixed effects as right-hand variables, while even-numbered columns also add demographic variables and baseline friend counts. In Panel A, we see that the estimates for outcome "Tried to Use" strongly increase when we move from one to two observed friends, then do not increase between two and three. In Panel B, the estimates for "Used" increase when moving from one to two and from two to three.

Table 2.5: Nepal Peer Effects Estimates (Months Pooled)

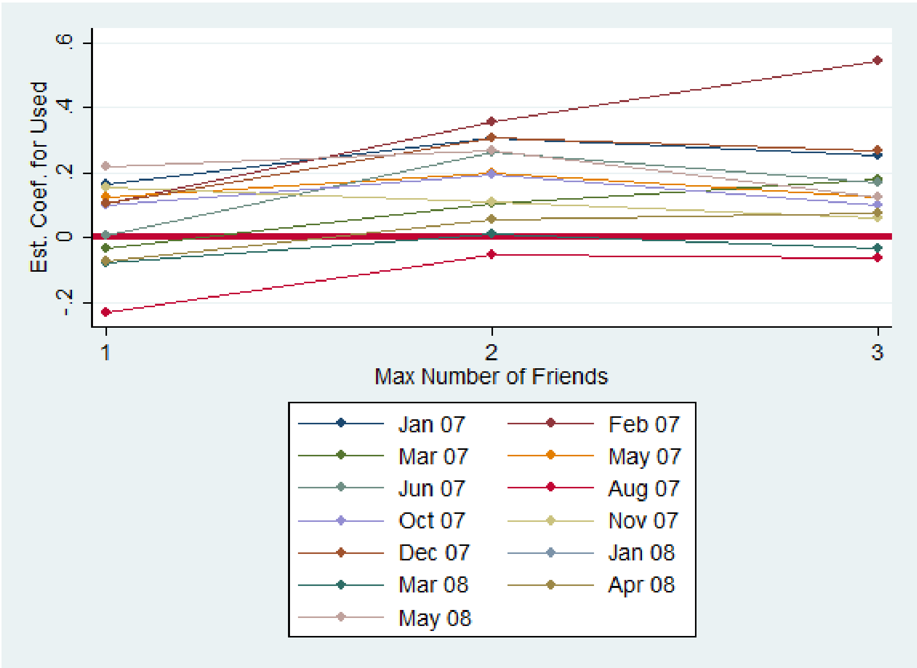
	(1)	(2)	(3)	(4)	(5)	(6)
Max Number of Friends	1	1	2	2	3	3
Panel A: Tried to Use						
Treated	0.059*	0.047	0.152***	0.178***	0.138**	0.153***
	(0.035)	(0.037)	(0.047)	(0.050)	(0.055)	(0.058)
Baseline Controls	NO	YES	NO	YES	NO	YES
R-squared	0.148	0.168	0.155	0.174	0.152	0.171
Panel B: Used						
Treated	0.129***	0.062*	0.274***	0.244***	0.335***	0.299***
	(0.036)	(0.035)	(0.049)	(0.047)	(0.056)	(0.054)
Baseline Controls	NO	YES	NO	YES	NO	YES
R-squared	0.256	0.349	0.271	0.364	0.274	0.367

Notes: N = 924 in all specifications. Standard errors in parentheses. All specifications include school fixed effects. *** p<0.01, ** p<0.05, * p<0.1. Sample restricted to observations with non-missing data. School fixed effects included in all specifications. Baseline controls include demographic variables and number of friends reported.

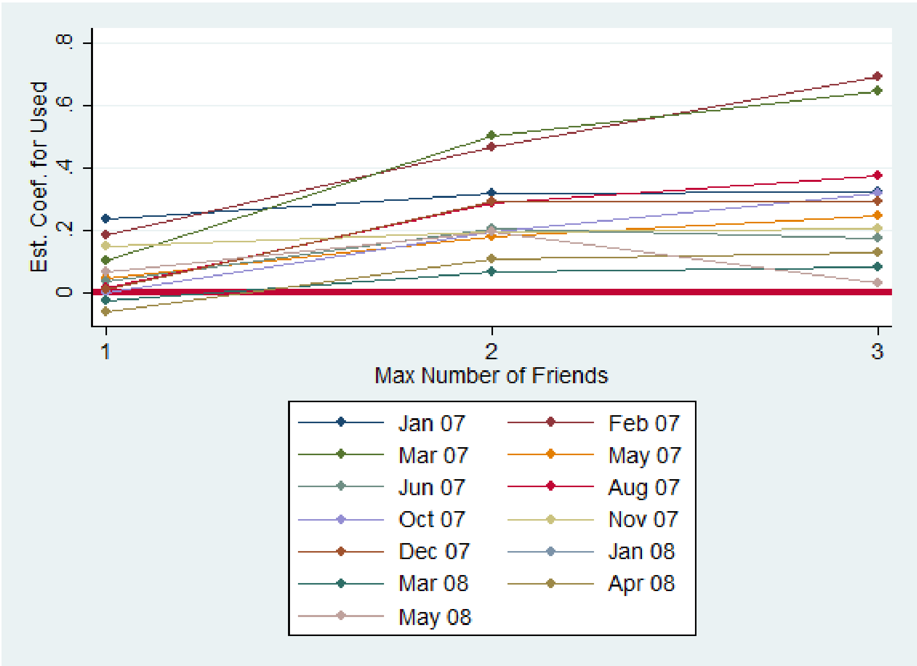
Next, in Figure 2.5, I present estimates of $\hat{\alpha}_2$ from Equation (2.12) run separately for each month in the sample. Similar to the pooled results, in Subfigure (a) we see a sharp increase in the coefficients when moving from 1 to 2 observee friends, and a smaller increase moving from 2 to 3. In Subfigure (b) there is a sharp upward trend when increasing observed friends both from 1 to 2 and from 2 to 3. These results provide further evidence of the extent to which censoring in network data may bias peer effects estimates.

Figure 2.5: Nepal Peer Effects Estimates (Disaggregated by Month)

(a) Coefficients for Tried to Use



(b) Coefficients for Used



2.6 Bias Correction Strategies

The prior sections showed that censoring may lead to bias. I show bias both analytically and by simulation in Sections 2.3 and 2.4, respectively. Then Section 2.5 presented results from two datasets suggesting that this bias is present in real datasets. Given these results, this section presents two strategies to deal with censoring-induced bias. Under some assumptions, both estimators lead to consistent estimates of the true parameters of the model. I present simulations of these bias-corrected estimators to show that they perform reasonably well when the data-generating process is known.

2.6.1 Uncensored Subsample

If the researcher has the ability to manipulate the data generating process, the best solution is to simply collect uncensored data. However, this may not always be practical due to cost. Accordingly, a middle ground may be to collect an uncensored subsample. Collecting this uncensored subsample allows for consistent estimation of the parameters of the model, either α or β.

Recall from Proposition 2.7, we know that $\operatorname{plim} \hat{\alpha}^{cens,k} = B_\alpha \alpha$. Accordingly, $\hat{\alpha}^{cens,k}$ is a consistent estimator of $B_\alpha \alpha$. Further, recall that

$$\mathbf{B}_\alpha = \begin{bmatrix} \mathbb{E}[\mathbf{x}'\mathbf{x}] & \mathbb{E}[\mathbf{x}'\mathbf{H_k}\mathbf{x}] \\ \mathbb{E}[\mathbf{x}'\mathbf{H_k}\mathbf{x}] & \mathbb{E}[\mathbf{x}'\mathbf{H_k}'\mathbf{H_k}\mathbf{x}] \end{bmatrix}^{-1} \begin{bmatrix} \mathbb{E}[\mathbf{x}'\mathbf{x}] & \mathbb{E}[\mathbf{x}'\mathbf{H_k}\mathbf{x}] \\ \mathbb{E}[\mathbf{x}'\mathbf{G}\mathbf{x}] & \mathbb{E}[\mathbf{x}'\mathbf{H_k}'\mathbf{G}\mathbf{x}] \end{bmatrix} \tag{2.13}$$

Accordingly, from the uncensored subsample, we can construct an empirical analogue of B_α as a plug-in estimator $\hat{B}_\alpha$. We can then construct $\hat{\alpha}_{BC} = \hat{B}_\alpha^{-1} \hat{\alpha}^{cens,k}$. With some regularity conditions, by the Slutzky Theorem, $\operatorname{plim} \hat{\alpha}_{BC} = \alpha$.

Similar logic applies for constructing a bias-corrected estimator $\hat{\beta}_{BC} = \hat{B}_\beta^{-1} \hat{\beta}^{cens,k}$, where $\operatorname{plim} \hat{\beta}_{BC} = \beta$.

2.6.2 Estimate on Uncensored Nodes

As an alternative bias-correction strategy, I slightly modify an estimator developed in Chandrasekhar and Lewis (2011). The idea here is to restrict estimation to nodes (individuals) that have no missing data. The authors of that paper were dealing with a slightly different context than the one at issue here. Their data were randomly sampled at the node level, so that the researchers observed all links for a randomly-sampled subset of the population. Measurement error in the network arose due to the fact that only 46% of the population was sampled. In that context, restricting estimation to the 46% who were sampled ensured that there was no measurement error on any regressors.

In the case of censoring, missingness is not random. Rather, we can only be sure individuals have no missing network links if they name fewer than the maximum allowable links on the survey. For this subset of the population, $\mathbf{H_k x} = \mathbf{Gx}$. Accordingly, there is no measurement error when we estimate $\hat{\alpha}$ restricted to these individuals.

Further, as pointed out by Chandrasekhar and Lewis (2011), if we estimate $\hat{\beta}$ restricted to those with no missing links, there will be measurement error on thier second-order links $\mathbf{G}^2\mathbf{x}$. However, this measurement error is uncorrelated with measurement error of their first-order links, which is uniformly zero. Crucially, for this subset of individuals, the measurement error of excluded instrument $\mathbf{H_k}^2\mathbf{x}$ is uncorrelated with (nonexistent) measurement error of the exogenous regressor $\mathbf{H_k y} = \mathbf{Gy}$. In sum, restricting estimation to those individuals for whom we know their entire network should produce consistent estimates of $\hat{\beta}$.

2.6.3 Simulations of Bias-Corrected Estimators

Table 2.6 presents median estimated coefficients for three estimators.[5]

1. Uncorr, the uncorrected estimator. These correspond to the estimates presented in Tables E.3 to E.5 and Figure 2.1.

[5] These estimates are quite noisy, especially at lower numbers of K, so means may be skewed by outliers.

2. US, the Uncensored Subsample method.

3. CL, the method derived in Chandrasekhar and Lewis (2011).

In all simulations, the unconditional probability of a link existing is 0.35, while I allow γ_1 to vary. From Panel A, we see that, as discussed above, whenever $\gamma_1 = 0$, there is no bias in censored estimates of $\hat{\alpha}_1$. Both bias-corrected estimators for $\hat{\alpha}_1$ perform much better than the uncorrected one whenever $\gamma_1 < 0$.

Panel B shows similar results for $\hat{\alpha}_2$, except that the uncorrected estimator is biased even when $\gamma_0 = 0$. Note that the bias-corrected estimators "US" and "CL" remain mostly constant for the entire range of K, while the uncorrected estimator trends strongly as K increases. That is, the medians of the bias-corrected estimators are not sensitive to the number of links observed, unlike the uncorrected estimator.

Estimates of $\hat{\beta}$ in Table 2.7 show similar patterns. In general, the uncorrected estimator is biased for small K, while both "US" and "CL" perform reasonably well at small K. Again note the exception that when $\gamma_1 = 0$, uncorrected $\hat{\beta}_2$ remains a consistent estimator of β_2 even in the presence of censoring.

From this section, we have learned that both the Uncensored Subsample and the CL estimator are consistent even in the presence of censoring. However, these simulations rely crucially upon a homogeneity assumption. When there is homogeneity on the marginal effects β, then the estimators may not return consistent estimates of the target parameters, which is the subject of the next section.

2.7 Heterogeneity and Random Coefficients

As shown above, the assumption of homogeneous marginal effects, along with exogeneity conditions, is sufficient for consistency of the bias-corrected estimators. However, the assumption of homogeneous marginal effects is a very strong one. In this section, I discuss the effects of relaxing this assumption, demonstrating that none of the estimators are consistent

Table 2.6: Comparison of Bias-Corrected $\hat{\alpha}$

	$\gamma_1 = 0$			$\gamma_1 = -1$			$\gamma_1 = -2$		
Estimator	Uncorr	US	CL	Uncorr	US	CL	Uncorr	US	CL
Panel A: $\hat{\alpha}_1$									
2	1.116	1.117	1.425	1.425	0.947	0.982	1.399	0.949	0.974
3	1.116	1.117	1.102	1.279	0.941	0.974	1.249	0.936	0.959
4	1.116	1.117	1.092	1.164	0.936	0.967	1.140	0.919	0.947
5	1.116	1.117	1.095	1.074	0.929	0.959	1.052	0.908	0.938
6	1.115	1.116	1.101	1.003	0.923	0.952	0.982	0.897	0.926
7	1.115	1.116	1.103	0.957	0.918	0.943	0.931	0.888	0.919
8	1.116	1.116	1.105	0.930	0.913	0.935	0.901	0.884	0.907
9	1.116	1.116	1.108	0.916	0.911	0.928	0.884	0.881	0.897
10	1.116	1.116	1.111	0.910	0.909	0.921	0.880	0.879	0.891
11	1.116	1.116	1.113	0.909	0.908	0.917	0.878	0.878	0.885
12	1.116	1.116	1.115	0.908	0.908	0.913	0.877	0.878	0.881
13	1.116	1.116	1.115	0.908	0.908	0.911	0.877	0.877	0.879
14	1.116	1.116	1.115	0.908	0.908	0.909	0.877	0.877	0.878
15+	1.116	1.116	1.116	0.908	0.908	0.908	0.877	0.877	0.878
Panel B: $\hat{\alpha}_2$									
2	0.520	1.639	1.420	0.811	1.987	1.920	1.342	2.237	2.316
3	0.762	1.625	1.280	1.099	1.999	1.900	1.584	2.261	2.286
4	0.994	1.619	1.310	1.357	2.007	1.909	1.785	2.278	2.273
5	1.201	1.625	1.360	1.584	2.014	1.912	1.957	2.297	2.271
6	1.375	1.623	1.412	1.770	2.029	1.922	2.101	2.312	2.273
7	1.499	1.626	1.467	1.904	2.043	1.945	2.211	2.321	2.282
8	1.573	1.626	1.506	1.986	2.049	1.967	2.280	2.330	2.292
9	1.611	1.628	1.542	2.029	2.053	1.986	2.318	2.337	2.303
10	1.626	1.630	1.572	2.047	2.055	2.007	2.334	2.340	2.312
11	1.629	1.631	1.597	2.053	2.056	2.023	2.340	2.341	2.321
12	1.631	1.631	1.613	2.055	2.056	2.037	2.342	2.342	2.331
13	1.631	1.631	1.622	2.056	2.056	2.045	2.342	2.342	2.336
14	1.631	1.631	1.628	2.056	2.056	2.051	2.342	2.342	2.339
15+	1.631	1.631	1.630	2.056	2.056	2.054	2.342	2.342	2.342

Table presents *median* simulated estimates. $Pr(link) = 0.35$ in all simulations. Standard deviations of simulated estimates in brackets. 1000 simulations performed with 200 schools of 25 students each (N = 5000). Uncensored subsample consists of 10% of the total sample (20 schools of 25 students each).

Table 2.7: Comparison of Bias-Corrected $\hat{\beta}$

	$\gamma_1 = 0$			$\gamma_1 = -1$			$\gamma_1 = -2$		
Estimator	Uncorr	US	CL	Uncorr	US	CL	Uncorr	US	CL
Panel A: $\hat{\beta}_1$									
2	0.726	0.605	0.950	0.693	0.625	0.597	0.979	0.598	0.583
3	0.746	0.602	0.600	0.701	0.605	0.607	0.789	0.595	0.595
4	0.734	0.602	0.589	0.708	0.596	0.602	0.725	0.602	0.597
5	0.704	0.600	0.597	0.701	0.595	0.599	0.708	0.597	0.599
6	0.671	0.600	0.594	0.682	0.598	0.598	0.691	0.598	0.598
7	0.641	0.599	0.598	0.656	0.597	0.600	0.669	0.597	0.598
8	0.619	0.598	0.597	0.632	0.596	0.597	0.645	0.596	0.599
9	0.607	0.599	0.599	0.614	0.599	0.601	0.622	0.598	0.598
10	0.602	0.599	0.598	0.606	0.598	0.600	0.609	0.598	0.599
11	0.599	0.598	0.599	0.601	0.599	0.600	0.602	0.598	0.599
12	0.599	0.599	0.600	0.599	0.598	0.601	0.599	0.598	0.597
13+	0.599	0.599	0.598	0.599	0.599	0.598	0.599	0.598	0.598
Panel B: $\hat{\beta}_2$									
2	0.971	0.996	1.256	1.361	0.990	0.997	1.360	0.992	1.000
3	0.967	0.997	1.002	1.253	0.992	1.002	1.262	0.997	1.000
4	0.969	0.998	0.998	1.174	0.992	1.001	1.185	0.993	1.000
5	0.976	0.998	0.997	1.113	0.997	1.000	1.124	0.996	1.000
6	0.983	0.999	0.998	1.067	0.998	1.001	1.074	0.998	1.000
7	0.990	1.000	0.998	1.035	0.999	1.000	1.040	0.997	0.999
8	0.995	0.999	0.998	1.016	1.000	1.000	1.018	0.997	1.000
9	0.998	0.999	0.999	1.007	1.000	0.999	1.005	0.997	0.999
10	0.998	0.999	0.999	1.003	1.001	1.000	1.000	0.998	0.999
11	0.999	0.999	0.999	1.001	1.001	1.000	0.999	0.998	0.998
12	0.999	0.999	0.999	1.001	1.001	1.001	0.999	0.998	0.998
13+	0.999	0.999	0.999	1.001	1.001	1.001	0.999	0.999	0.998
Panel C: $\hat{\beta}_3$									
2	-0.409	0.495	0.820	-0.608	0.485	0.512	-1.240	0.532	0.539
3	-0.288	0.498	0.451	-0.414	0.508	0.482	-0.554	0.543	0.514
4	-0.130	0.495	0.480	-0.241	0.526	0.494	-0.240	0.522	0.508
5	0.044	0.502	0.501	-0.062	0.522	0.503	-0.077	0.526	0.502
6	0.211	0.504	0.503	0.112	0.511	0.497	0.073	0.517	0.508
7	0.344	0.502	0.497	0.261	0.514	0.496	0.208	0.512	0.514
8	0.430	0.503	0.497	0.374	0.509	0.505	0.330	0.516	0.506
9	0.474	0.501	0.499	0.449	0.507	0.502	0.415	0.510	0.504
10	0.493	0.501	0.499	0.483	0.506	0.500	0.474	0.514	0.501
11	0.499	0.500	0.502	0.500	0.505	0.502	0.496	0.510	0.505
12	0.500	0.501	0.502	0.504	0.507	0.502	0.504	0.509	0.511
13+	0.500	0.501	0.502	0.505	0.505	0.504	0.508	0.508	0.508

Table presents *median* simulated estimates. $Pr(link) = 0.35$ in all simulations. Standard deviations of simulated estimates in brackets. 1000 simulations performed with 200 schools of 25 students each (N = 5000). Uncensored subsample consists of 10% of the total sample (20 schools of 25 students each).

even in the absence of censoring.

2.7.1 Set-up and Defining the Target Parameter

Suppose that we allow heterogeneity on β. That is, for each student i in school s, there exists a marginal effect β_{is} that may be non-constant. Accordingly, a generalization of Equation (2.2) is Equation (2.14).

$$y_{is} = \beta_0 + \bar{y}_{is}\beta_{1is} + x_{is}\beta_{2is} + \bar{x}_{is}\beta_{3is} + \epsilon_{is} \tag{2.14}$$

Since β_{is} is now drawn from some distribution, we must define the parameter of interest. For the sake of simplicity, I assume that $\beta_{is} \in \mathbb{B}$, a compact subset of $\mathbb{R}^3$. Accordingly, the first moment exists. While other features of this distribution may be of interest, for purposes here I assume that the parameter of interest is $\mathbb{E}[\beta_{is}] = \bar{\beta}$, defined in Wooldridge (2010) as the *average partial effect* (see also Heckman and Vytlacil, 1998; Wooldridge, 2003). This is a generalization of the notion of *average treatment effect* that takes prominence in much of the treatment effects literature.

Assumption 2.3. $\mathbb{E}[\epsilon_{is}|\beta_{js}] = 0 \; \forall \; j$

In this entire section, I assume exogeneity of the unobserved β_{is}. This assumption is formalized in Assumption 2.3. Essentially, this requires ϵ_{is} is (mean) independent of heterogeneity in the effect of regressors on y_{is}. This mean independence assumption is standard in the literature on correlated random effects (Wooldridge, 1997; Heckman and Vytlacil, 1998; Wooldridge, 2003). Given Assumption 2.3, primary interest lies in consistently estimating $\bar{\beta}$ and $\hat{\alpha}$, which is defined below.

2.7.2 To What Do the Estimators Converge in the Presence of Heterogeneity

First, I make an assumption on the data-generating process. Unobserved heterogeneity in the marginal effect β_{is} needs to be mean independent of observable $\mathbf{x}$ and the network

$\mathbf{G}$. This is an admittedly strong assumption that rules out dependence between unobserved β_{is} and x_{is}. the literature on correlated random effects discusses identification when this assumption fails (See Wooldridge, 1997; Heckman and Vytlacil, 1998; Wooldridge, 2003).

Assumption 2.4. $\mathbb{E}[\beta_{is}|\mathbf{x}, \mathbf{G}]$ *(Uncorrelated Random Effects)*

Given Assumptions 2.3 and 2.4,

$$\text{plim}\,\hat{\alpha} = \begin{bmatrix} \bar{\beta}_2 \\ \bar{\beta}_3 + \mathbb{E}[\beta_{1is}\beta_{3is}] \end{bmatrix} + (\sum_{k=1}^{\infty} \mathbb{E}[(\beta_{3is} + \beta_{1is}\beta_{2is})\beta_{1is}^k]\,(\mathbb{E}[[\mathbf{x}, \mathbf{Gx}]'[\mathbf{x}, \mathbf{Gx}]])^{-1} \begin{bmatrix} \mathbb{E}[\mathbf{x}'\mathbf{G}^{k+1}\mathbf{x}] \\ \mathbb{E}[\mathbf{x}'\mathbf{G}'\mathbf{G}^{k+1}\mathbf{x}] \end{bmatrix} \tag{2.15}$$

Importantly, β_{is} being independent of the network-formation process is a crucial prerequisite. I note that $\text{plim}\,\hat{\alpha}$ is a very complicated expression that again depends on the relationship between observed characteristics $\mathbf{x}$, the network $\mathbf{G}$, and unobserved heterogeneity β_{is}.

Further, and crucially, $\hat{\beta}$ is not a consistent estimator of $\bar{\beta}$. This is due to a variation on the reflection problem. To see this, define the following matrices:

$$\mathbf{W} = \begin{bmatrix} \mathbf{Gy} & \mathbf{x} & \mathbf{Gx} \end{bmatrix}$$
$$\mathbf{Z} = \begin{bmatrix} \mathbf{G}^2\mathbf{x} & \mathbf{x} & \mathbf{Gx} \end{bmatrix}$$

From this, $w_{is}, z_{is} \in \mathbb{R}^3$ are row vectors that denote the corresponding elements of matrices $\mathbf{W}$ and $\mathbf{Z}$. Accordingly, with Assumptions 2.3 and 2.4,

$$\text{plim}\,\hat{\beta}^{uncens} = \mathbb{E}[z_{is}' w_{is}]^{-1}\mathbb{E}[z_{is}' w_{is}\beta_{is}] \tag{2.16}$$

From this, it is clear that $\mathbb{E}[\beta_{is}|w_{is}] = \bar{\beta}$ is sufficient for consistency. However, whenever β_{is} is non-constant, this condition will not hold. Assuming $\beta_{is} > 0$, then higher β_{is} implies higher y_{is}. Since i affects his peers, this implies higher $\bar{y}_{is}$. Accordingly, *even without censoring*, when β_{is} allows for heterogeneity, the estimator $\hat{\beta}$ is not in general a consistent estimator of $\bar{\beta}$.

This result is closely related to the standard reflection problem identified by Manski (1993). He identified the correlation between $\bar{y}_{is}$ and unobserved ϵ_{is} as leading to necessary endogeneity in the system. Correlation between $\bar{y}_{is}$ and β_{is} leads to inconsistency in a similar manner.

This is not an if and only if condition, however. Rather, this result only states that a particular and widely-known condition for consistency does not hold for this particular IV estimator. It is possible that other reasonable conditions may be found, but the simulations provided hereafter suggest that the estimator is inconsistent for a wide variety of parameter values.

2.7.3 When Data Is Censored

The prior discussed the limits of the estimators $\hat{\alpha}$ and $\hat{\beta}$ when there is no censoring. From that, we saw that $\hat{\alpha}$ converges to a complicated expression while $\hat{\beta}$ in general does not converge to $\bar{\beta}$.

Define an individual's friend count as $F_{is} \leq M$. Similar to the uncensored result in Equation 2.15, when data is censored at $K \leq M$,

$$\operatorname{plim} \hat{\alpha}^{K,cens} = \begin{bmatrix} \bar{\beta}_2 \\ \bar{\beta}_3 + \mathbb{E}[\beta_{1is}\beta_{3is}|F_{is} \leq K] \end{bmatrix} + (\sum_{k=1}^{\infty} \mathbb{E}[(\beta_{3is} + \beta_{1is}\beta_{2is})\beta_{1is}^k|F_{is} \leq K]$$
$$(\mathbb{E}[[\mathbf{x}, \mathbf{Gx}]'[\mathbf{x}, \mathbf{Gx}]|F_{is} \leq K])^{-1} \begin{bmatrix} \mathbb{E}[\mathbf{x'G}^{k+1}\mathbf{x}|F_{is} \leq K] \\ \mathbb{E}[\mathbf{x'G'G}^{k+1}\mathbf{x}|F_{is} \leq K]) \end{bmatrix} \tag{2.17}$$

Accordingly, if we are interested in $\operatorname{plim} \hat{\alpha}$ (uncensored), a sufficient condition for $\operatorname{plim} \hat{\alpha}^{K,cens} = \operatorname{plim} \hat{\alpha}$ is that the conditional expectations are equal to the unconditional ones. Importantly, this requires independence of unobserved heterogeneity β_{is} from the number of friends F_{is}. This will clearly fail if, for example, those with higher unobserved ability (captured by β_{is})

also tend to more more friends (captured by F_{is}).

Analogous to the uncensored result in Equation 2.16, with data censored at $K \leq M$,

$$\text{plim}\,\hat{\beta}^{cens,K} = \mathbb{E}[z_{is}' w_{is} | F_{is} < K]^{-1} \mathbb{E}[z_{is}' w_{is} \beta_{is} | F_{is} < K] \tag{2.18}$$

Clearly, plim $\hat{\beta}^{cens,K}$ depends on the censoring rule K whenever β_{is} and F_{is} are correlated, such as when unobserved ability is correlated with number of friends.

2.7.4 Simulations with Heterogeneity

For purposes of simplicity, I assume a simple data-generating process. Links are formed according to the same rule as in Section 2.4. However, I introduce heterogeneity in β_{is} that depends on degree in the following manner: β_{is} takes on two values, with switching at the median of the degree distribution. The set-up is described in Table 2.8.

Table 2.8: Values of β_{is} for Simulations with Heterogeneity

Parameter	Below Median	Above Median	Exp
β_{1is}	0.6	0.2	0.4
β_{2is}	1	0.5	0.75
β_{3is}	0.5	0.25	0.375

Table 2.9 presents medians of simulations of $\hat{\alpha}$. For each level of K and γ_1, I present the raw uncorrected medians as well as the medians of the "US" and "CL" estimators. From Panel A, we see that, as before, when $\gamma_1 = 0$, censoring leads to no bias in $\hat{\alpha}$, while CL is in biased. This is due to the fact that, for small levels of K, the CL estimator only uses those who are uncensored and thus have the β_{is} for the lower half of the degree distribution.

Table 2.10 presents analogous results for $\hat{\beta}$. From this we see that, as discussed above, even without censoring $K = 15+$, none of the estimators is consistent for $\mathbb{E}[\beta_{is}] = (0.4, 0.75, 0.375)'$, except in the single case of $\hat{\beta}_2$ when $\gamma_1 = 0$ (and thus $\mathbf{x}$ and $\mathbf{Gx}$ are independent).

Table 2.9: Comparison of Bias-Corrected $\hat{\alpha}$ (with Heterogeneity)

	$\gamma_1 = 0$			$\gamma_1 = -1$			$\gamma_1 = -2$		
Estimator	Uncorr	US	CL	Uncorr	US	CL	Uncorr	US	CL
Panel A: $\hat{\alpha}_1$									
2	0.791	0.792	2.116	1.314	1.058	0.968	1.355	1.065	0.952
3	0.791	0.792	1.084	1.235	1.059	0.954	1.276	1.080	0.920
4	0.791	0.792	1.055	1.170	1.052	0.944	1.209	1.076	0.896
5	0.791	0.791	1.064	1.113	1.039	0.942	1.141	1.056	0.878
6	0.791	0.791	1.062	1.063	1.021	0.943	1.075	1.025	0.865
7	0.791	0.791	1.063	1.020	1.002	0.951	1.012	0.987	0.862
8	0.791	0.791	1.064	0.988	0.980	0.959	0.959	0.950	0.861
9	0.791	0.791	1.064	0.968	0.964	0.962	0.922	0.920	0.867
10	0.791	0.791	0.935	0.958	0.958	0.955	0.907	0.906	0.872
11	0.791	0.791	0.863	0.955	0.955	0.953	0.901	0.901	0.879
12	0.791	0.791	0.824	0.954	0.954	0.952	0.899	0.899	0.886
13	0.791	0.791	0.806	0.954	0.954	0.952	0.898	0.898	0.891
14	0.791	0.791	0.796	0.954	0.954	0.952	0.898	0.898	0.894
15+	0.791	0.791	0.792	0.954	0.954	0.953	0.898	0.898	0.896
Panel B: $\hat{\alpha}_2$									
2	0.265	0.856	1.831	0.461	1.085	1.872	0.879	1.450	2.282
3	0.390	0.859	1.002	0.618	1.080	1.831	1.012	1.426	2.247
4	0.510	0.857	1.009	0.764	1.097	1.806	1.135	1.426	2.225
5	0.617	0.852	1.027	0.902	1.122	1.781	1.262	1.460	2.206
6	0.707	0.849	1.033	1.025	1.159	1.753	1.385	1.508	2.184
7	0.773	0.848	1.047	1.129	1.199	1.717	1.497	1.565	2.155
8	0.812	0.845	1.063	1.207	1.236	1.683	1.592	1.619	2.124
9	0.833	0.844	1.075	1.257	1.269	1.635	1.653	1.664	2.053
10	0.840	0.843	0.968	1.280	1.284	1.500	1.681	1.685	1.923
11	0.842	0.843	0.906	1.288	1.290	1.409	1.692	1.693	1.831
12	0.843	0.843	0.871	1.290	1.290	1.353	1.696	1.697	1.769
13	0.843	0.843	0.855	1.291	1.291	1.323	1.697	1.697	1.734
14	0.843	0.843	0.848	1.291	1.291	1.304	1.698	1.698	1.714
15+	0.843	0.843	0.845	1.291	1.291	1.297	1.698	1.698	1.703

Table presents *median* simulated estimates. $Pr(link) = 0.35$ in all simulations. Standard deviations of simulated estimates in brackets. 1000 simulations performed with 200 schools of 25 students each (N = 5000).

Table 2.10: Comparison of Bias-Corrected $\hat{\beta}$ (with Heterogeneity)

	$\gamma_1 = 0$			$\gamma_1 = -1$			$\gamma_1 = -2$		
Estimator	Uncorr	US	CL	Uncorr	US	CL	Uncorr	US	CL
Panel A: $\hat{\beta}_1$									
2	0.547	0.216	1.391	0.564	0.281	0.644	1.110	0.556	0.592
3	0.591	0.286	0.633	0.551	0.339	0.596	0.724	0.427	0.606
4	0.569	0.308	0.594	0.542	0.336	0.592	0.618	0.432	0.600
5	0.516	0.321	0.609	0.517	0.338	0.596	0.623	0.483	0.601
6	0.452	0.321	0.581	0.466	0.320	0.602	0.635	0.536	0.599
7	0.395	0.321	0.601	0.410	0.307	0.601	0.629	0.563	0.601
8	0.358	0.323	0.601	0.344	0.291	0.603	0.611	0.571	0.599
9	0.343	0.329	0.601	0.298	0.274	0.597	0.589	0.568	0.603
10	0.335	0.331	0.483	0.274	0.269	0.517	0.578	0.570	0.610
11	0.333	0.330	0.407	0.263	0.261	0.448	0.572	0.568	0.608
12	0.333	0.332	0.370	0.259	0.259	0.382	0.570	0.569	0.596
13+	0.333	0.333	0.347	0.259	0.258	0.323	0.569	0.569	0.587
Panel B: $\hat{\beta}_2$									
2	0.735	0.775	1.973	1.274	1.059	1.004	1.321	1.097	0.998
3	0.732	0.771	0.975	1.210	1.063	1.004	1.275	1.107	0.998
4	0.735	0.766	0.999	1.157	1.057	1.002	1.227	1.109	0.997
5	0.740	0.764	1.004	1.109	1.045	1.003	1.179	1.102	0.999
6	0.746	0.762	1.001	1.065	1.027	1.002	1.133	1.087	0.999
7	0.752	0.762	1	1.027	1.008	1.003	1.085	1.059	0.997
8	0.757	0.761	1.001	0.997	0.988	1.003	1.042	1.031	0.997
9	0.759	0.761	1	0.975	0.972	0.999	1.010	1.004	0.996
10	0.760	0.760	0.888	0.966	0.965	0.980	0.994	0.992	0.994
11	0.760	0.760	0.823	0.963	0.962	0.970	0.988	0.987	0.991
12	0.760	0.760	0.790	0.962	0.961	0.965	0.986	0.985	0.989
13+	0.760	0.760	0.773	0.961	0.961	0.963	0.985	0.985	0.988
Panel C: $\hat{\beta}_3$									
2	-0.171	0.744	1.680	-0.431	0.609	0.385	-1.470	0.253	0.515
3	-0.110	0.636	0.475	-0.276	0.520	0.508	-0.540	0.488	0.489
4	0.003	0.583	0.485	-0.128	0.525	0.516	-0.218	0.450	0.505
5	0.137	0.561	0.498	0.027	0.546	0.507	-0.131	0.354	0.501
6	0.277	0.545	0.514	0.226	0.601	0.499	-0.071	0.257	0.507
7	0.388	0.536	0.504	0.426	0.663	0.489	0.031	0.234	0.504
8	0.462	0.531	0.502	0.604	0.723	0.485	0.142	0.255	0.502
9	0.498	0.526	0.501	0.738	0.792	0.494	0.250	0.307	0.468
10	0.512	0.520	0.504	0.801	0.820	0.532	0.297	0.322	0.363
11	0.518	0.520	0.510	0.826	0.832	0.605	0.322	0.330	0.316
12	0.519	0.520	0.507	0.832	0.832	0.678	0.330	0.333	0.308
13+	0.520	0.520	0.515	0.836	0.836	0.746	0.334	0.336	0.308

Table presents *median* simulated estimates. $Pr(link) = 0.35$ in all simulations. Standard deviations of simulated estimates in brackets. 1000 simulations performed with 200 schools of 25 students each (N = 5000). $\mathbb{E}[\beta_{1is}] = 0.4$, $\mathbb{E}[\beta_{2is}] = 0.75$, $\mathbb{E}[\beta_{3is}] = 0.25$ in all simulations.

2.8 A Note on Graphical Reconstruction

An additional method is more complicated and relies upon graphical reconstruction. This is the primary method developed in Chandrasekhar and Lewis (2011). The idea is to use the uncensored observations to estimate a model that can then be used to impute values on the censored ones. As the authors of that paper note, this requires placing substantial structure on the problem. Any method of graphical reconstruction requires first specifying a model for how the network is formed.Recent papers have extended this approach using statistical random graph models for graphical reconstruction (Chandrasekhar and Jackson, 2014; Williams, 2016).

The method developed by Chandrasekhar and Lewis (2011) is closely related to methods for multiple imputation (See Rubin, 1976; Cameron and Trivedi, 2005). The validity of these methods requires strong assumptions that are likely to fail in the case of censored data. Most importantly, the assumption that data is *missing conditionally at random* (MCAR) does not hold in this setting. Consider the simple model of network formation used to generate the simulations in Section 2.4. Links are listed in order of utility U_{ijs}, such that

$$l_{ijs} = \mathbf{1}\{\gamma_0 + |X_{is} - X_{js}|\gamma_1 + u_{ijs} > 0\} = \mathbf{1}\{U_{ijs} > 0\}$$

The MCAR assumption requires u_{ijs} be distributed the same for observed and unobserved (censored) links. This is clearly not going to be true. Rather, even conditional on $|X_{is} - X_{js}|$, u_{ijs} will tend to be higher for observed links than unobserved ones. Even if we correctly specify the functional form of the data-generating process, the fact that data are not missing conditionally at random implies that we will get biased estimates of the parameters $\hat{\gamma}$, leading to erroneous graphical reconstruction.

2.9 Conclusion

As economists have moved beyond a classroom model of peer effects, the collecting of accurate network data has become more and more important. However, many efforts to collect such data censor the number of links that may be listed. This paper investigates the implications of censoring for estimates of peer effects and to provides strategies to overcome these limitations. In doing so, I make two primary contributions to the literature.

First, I document the potential for bias in estimates of linear-in-means models that are estimated with censored network data. I do this first analytically then by simulation. Then I estimate these same parameters while varying the number of links observed in two different datasets. These results all suggest that censoring may crucially bias estimates.

Second, I suggest two strategies for recovering consistent estimates of the target parameters. First, censored estimates can be corrected analytically by estimating a bias-correction term in an uncensored subsample. Next, I adapt an estimator derived in Chandrasekhar and Lewis (2011) that estimates the model restricted to uncensored nodes. I show that both perform well in simulations, even when only a small number of links are observed. Finally, I discuss when the bias correction strategies may fail, notably in the case of heterogeneous effects. I also provide a brief discussion for why graphical reconstruction, as derived in Chandrasekhar and Lewis (2011) is unlikely to lead to consistent estimates in this context.

Through the analysis of this paper, I aim to highlight the real potential for bias due to censored network data to researchers collecting data. Attention to these issues may lead to better data-collection methods and to more careful consideration of the sensitivity of estimates to such methods.

CHAPTER 3

Network Partitioning and Social Exclusion under Different Selection Regimes

Joint work with

Clara Delavallade, International Food Policy Research Institute

Rebecca Thornton, University of Illinois at Urbana-Champaign

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.

3.1 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; Christakis and Fowler, 2007; Burgess and Umana-Aponte, 2011) or understanding how networks are formed (Burgess, Sanderson and Aponte, 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, Sanderson and Aponte, 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). 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 (DeGiorgi, Pelllizzari and Radaelli, 2010; Sacerdote, 2001), or when interventions are unlikely to affect network structure (Ngatia, 2015). 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.

In addition, there has been very little research accounting for changing network structure. In a recent paper, Comola and Prina (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 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 the next subsection, we present the program background. The experimental design and data are described in Section 3.2. Network results are shown and discussed in Section 3.3, program effects on self-confidence and aspirations in Section 3.4. We conclude in Section 3.5.

3.1.1 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, 2013). Limited belief in one's own ability and self-efficacy translates into low aspirations and educational goals (Bandura et al., 2001) 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, Levy and Linden, 2013), female role models (Nguyen, 2008),(Beaman et al., 2009; Beaman, Duflo and adn Petia Topalova, 2012) (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.

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: 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.

3.2 Research Design

3.2.1 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:

1. OR friends: either i or j identifies the other as as a friend (L_{ijt}^{OR})

2. 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.

3.2.1.1 Baseline Data

Table 3.1 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

Table 3.1: Baseline Sample

	Girls (N=1414) (1)	Boys (N=1196) (2)	Difference (3)	P-Value of Test of Equality (4)
Standard	6.931	7.019	-0.088	0.086
	(0.001)	(0.001)	(0.049)	
Age	12.325	12.558	-0.232	0.021
	(0.004)	(0.008)	(0.095)	
Scheduled Caste	0.255	0.327	-0.072	0.030
	(0.001)	(0.002)	(0.032)	
Scheduled Tribe	0.123	0.262	-0.139	0.000
	(0.001)	(0.002)	(0.029)	
Other Backward Caste	0.445	0.284	0.161	0.000
	(0.002)	(0.001)	(0.036)	
Enrolled Previous Year	0.840	0.846	-0.006	0.833
	(0.001)	(0.001)	(0.029)	
Owns TV	0.868	0.779	0.089	0.013
	(0.001)	(0.001)	(0.034)	
Father Attended School	0.831	0.694	0.137	0.000
	(0.001)	(0.002)	(0.035)	
Mother Attended School	0.562	0.394	0.168	0.000
	(0.001)	(0.002)	(0.040)	
Education Index	-0.180	0.221	-0.401	0.001
	(0.010)	(0.007)	(0.106)	
Career Index	-0.129	0.151	-0.281	0.012
	(0.006)	(0.008)	(0.104)	
Self-Confidence Index	-0.009	0.008	-0.017	0.856
	(0.008)	(0.002)	(0.094)	
Gender Roles Index	0.112	-0.135	0.248	0.020
	(0.011)	(0.008)	(0.101)	
Number of Friends (OR)	15.761			
	(2.353)			
Number of Friends (AND)	7.815			
	(0.640)			

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. They are AND friends if they both name each other as friends.

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.

Table 3.1 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 with p=0.001 and 0.281 standard deviations with p=0.012, respectively). 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 earlier.

3.2.1.2 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.

3.2.2 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.

3.2.2.1 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:

$$BaseVar_{is} = \beta_0 + \beta_1 T1_s + \beta_2 T2_s + u_{is} \tag{3.1}$$

where $T1_s$ and $T2_s$ indicate being assigned to T1 or T2 schools. We estimate these regression separately for girls and boys and cluster standard errors by school . Table 3.2 presents the estimated regression coefficients β_0, β_1, and β_2, and the p-values indicating the statistical significance of the joint test of β_1 and β_2. Note that the null that means are the same across the three treatment groups is not rejected for girls for any baseline characteristic, as indicated by the p-values presented in Columns (4). Samples are also generally balanced among our baseline measures of empowerment and networks. Table 3.3 presents similar results for boys. However, with the small number of clusters, balance may be an issue: T2 girls have significangly lower baseline career aspirations. Relatedly, as shown in Table 3.2, 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.

Table 3.2: Baseline Balance – School-Level Randomization (Girls)

	Control (1)	Reg Coef on T1 (2)	Reg Coef on T2 (3)	P-Value of Joint Test (4)
Standard	6.934	0.069	-0.073	0.193
	(0.016)	(0.047)	(0.070)	
Age	12.458	-0.217	-0.212	0.281
	(0.112)	(0.152)	(0.146)	
Scheduled Caste	0.293	-0.097	-0.030	0.550
	(0.076)	(0.095)	(0.100)	
Scheduled Tribe	0.070	0.068	0.103	0.218
	(0.026)	(0.045)	(0.086)	
Other Backward Caste	0.430	0.016	0.032	0.957
	(0.083)	(0.099)	(0.108)	
Enrolled Previous Year	0.829	0.034	0.002	0.860
	(0.050)	(0.069)	(0.078)	
Owns TV	0.904	-0.080	-0.035	0.335
	(0.022)	(0.056)	(0.046)	
Father Attended School	0.848	0.005	-0.059	0.461
	(0.040)	(0.058)	(0.056)	
Mother Attended School	0.561	0.065	-0.065	0.310
	(0.055)	(0.076)	(0.085)	
Education Index	-0.212	-0.017	0.118	0.785
	(0.173)	(0.262)	(0.213)	
Career Index	-0.033	0.060	-0.362	0.034
	(0.119)	(0.132)	(0.184)	
Self-Confidence Index	0.075	-0.317	0.053	0.107
	(0.148)	(0.206)	(0.173)	
Gender Roles Index	0.194	-0.077	-0.178	0.770
	(0.116)	(0.221)	(0.254)	
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)	

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. They are AND friends if they both name each other as friends.

Table 3.3: Baseline Balance – School-Level Randomization (Boys)

	Control (1)	Reg Coef on T1 (2)	Reg Coef on T2 (3)	P-Value of Joint Test (4)
Standard	6.997	-0.030	0.114	0.114
	(0.053)	(0.059)	(0.081)	
Age	12.713	-0.339	-0.081	0.170
	(0.160)	(0.197)	(0.205)	
Scheduled Caste	0.399	-0.136	-0.068	0.406
	(0.066)	(0.100)	(0.115)	
Scheduled Tribe	0.201	0.062	0.125	0.524
	(0.050)	(0.070)	(0.133)	
Other Backward Caste	0.282	0.007	-0.002	0.991
	(0.050)	(0.070)	(0.081)	
Enrolled Previous Year	0.843	-0.002	0.016	0.931
	(0.041)	(0.058)	(0.055)	
Owns TV	0.788	-0.002	-0.031	0.964
	(0.048)	(0.062)	(0.115)	
Father Attended School	0.648	0.077	0.068	0.748
	(0.096)	(0.101)	(0.107)	
Mother Attended School	0.337	0.136	0.028	0.183
	(0.070)	(0.080)	(0.107)	
Education Index	0.315	-0.106	-0.195	0.681
	(0.134)	(0.185)	(0.228)	
Career Index	0.353	-0.305	-0.311	0.311
	(0.186)	(0.210)	(0.204)	
Self-Confidence Index	0.051	-0.093	-0.023	0.655
	(0.076)	(0.105)	(0.106)	
Gender Roles Index	0.061	-0.203	-0.440	0.235
	(0.167)	(0.188)	(0.251)	

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. They are AND friends if they both name each other as friends.

3.2.2.2 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:

$$BaseVar_{is} = \beta_0 + \beta_1 Selected_{is} + u_{is} \tag{3.2}$$

where $Selected_{is}$ indicates being randomly selected for parliament participation. We cluster standard errors by school. Table 3.4 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.

3.2.3 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 2,773 students, of whom 26,55 (95.7%) are found in administrative enrollment records.

We formally test for differential attrition by random assignment with Equations (3.1) and (3.2) in Table 3.5, using indicators of survey completion as dependent variables. Completion of a survey (Baseline Network, Endline Qnr or Endline Network) was not significantly associated with random assignment of intervention arms by school.

Table 3.4: Baseline Balance – Individual Randomization to Program among T2 Girls

	Not Selected (1)	Reg Coef on Selected (2)	P-Value of Joint Test (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 Backward 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)	

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. They are AND friends if they both name each other as friends.

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 3.5, 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:

$$
\begin{aligned}
Complete_{is} &= \gamma_0 + \gamma_1 T1_s + \gamma_2 T2_s + \gamma_3 BaseVar_{is} \\
&\quad + \gamma_4 T1_s \times BaseVar_{is} + \gamma_5 T2_s \times BaseVar_{is} + u_{is} \qquad (3.3) \\
Complete_{is} &= \beta_0 + \beta_1 Selected_{is} + \beta_2 BaseVar_{is} + \beta_3 Selected_{is} \times BaseVar_{is} + u_{is} \qquad (3.4)
\end{aligned}
$$

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 (3.3) (Appendix FTable F.1, Panel A), and analogously test for the significance of β_3 in equation (3.4) (Appendix F Table F.1, Panel B), which would indicate differential attrition. We present estimates of Equation (3.3) for girls and boys, while restricting estimation of Equation (3.4) 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.

3.3 Results

3.3.1 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

Table 3.5: Survey Attrition

Panel A: School-Level Randomization (Girls)

	Control (1)	Reg Coef on T1 (2)	Reg Coef on T2 (3)	P-Value of Joint Test (4)
Baseline Survey	0.688	0.099	0.072	0.240
	(0.042)	(0.059)	(0.057)	
Baseline Network Survey	0.710	-0.015	0.033	0.573
	(0.033)	(0.044)	(0.048)	
Endline Survey	0.714	0.071	0.031	0.276
	(0.030)	(0.043)	(0.048)	
Endline Network Survey	0.712	0.063	0.033	0.391
	(0.031)	(0.046)	(0.042)	

Panel B: School-Level Randomization (Boys)

	Control (1)	Reg Coef on T1 (2)	Reg Coef on T2 (3)	P-Value of Joint Test (4)
Baseline Survey	0.784	-0.011	-0.140	0.041
	(0.050)	(0.062)	(0.064)	
Baseline Network Survey	0.661	0.075	-0.012	0.182
	(0.059)	(0.066)	(0.069)	
Endline Survey	0.753	0.049	-0.054	0.172
	(0.033)	(0.045)	(0.056)	
Endline Network Survey	0.779	0.016	-0.077	0.397
	(0.023)	(0.039)	(0.064)	

Panel C: Randomization Among T2 Girls

	Not Selected (1)	Reg Coef on Selected (2)	P-Value of Joint Test (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)	

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. They are AND friends if they both name each other as friends.

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 ($linktype \in \{OR, AND\}$):

$$L_{ijs0}^{linktype} = \alpha_0 + \alpha_1 |BaseVar_{is} - BaseVar_{js}| + \epsilon_{ijs0} \tag{3.5}$$

$BaseVar_{is}$ 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 3.6 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

Table 3.6: Baseline Endogenously-Formed OR and AND Networks

Network Definition	OR (1)	AND (2)	Observations (3)
Standard	-0.102	-0.154	27418
	(0.012)	(0.018)	
Age	-0.019	-0.031	27298
	(0.007)	(0.010)	
Scheduled Caste	0.001	0.000	27210
	(0.020)	(0.030)	
Scheduled Tribe	-0.067	-0.057	27210
	(0.061)	(0.042)	
Other Backward Caste	-0.010	-0.014	27210
	(0.015)	(0.019)	
Enrolled Previous Year	-0.092	-0.117	24636
	(0.033)	(0.040)	
Family Owns TV	-0.040	-0.048	14428
	(0.032)	(0.060)	
Father Attended School	-0.005	-0.010	13774
	(0.024)	(0.038)	
Mother Attended School	-0.019	-0.033	13820
	(0.024)	(0.027)	
Education Index	-0.005	-0.027	13178
	(0.012)	(0.017)	
Career Index	0.004	-0.013	18070
	(0.014)	(0.018)	
Self-Confidence Index	-0.009	-0.023	18868
	(0.008)	(0.012)	
Gender Roles Index	-0.031	-0.049	18424
	(0.016)	(0.017)	
Number of Friends (OR)	-0.004	-0.008	27418
	(0.001)	(0.002)	
Number of Friends (AND)	0.002	-0.008	27418
	(0.001)	(0.003)	

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 network definition.
Reported values are regression coefficient on distance between students' values for each variable.

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.

3.3.2 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 3.7 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:

$$L_{ijs0}^{linktype} = \gamma_0 + \gamma_1 E_{ijs}^{OR} + \gamma_3 E_{ijs}^{AND} + \epsilon_{ijs0} \tag{3.6}$$

Table 3.7: Baseline Characteristics of Girls Elected and Not Elected

	Elected (N=374) (1)	Not Elected (N=1040) (2)	Difference (3)	P-Value of Joint Test (4)
Standard	7.112	6.866	0.246	0.000
	(0.058)	(0.029)	(0.059)	
Age	12.497	12.263	0.234	0.030
	(0.101)	(0.062)	(0.102)	
Scheduled Caste	0.273	0.248	0.025	0.569
	(0.058)	(0.036)	(0.043)	
Scheduled Tribe	0.094	0.134	-0.040	0.127
	(0.038)	(0.032)	(0.025)	
Other Backward Caste	0.422	0.453	-0.030	0.546
	(0.056)	(0.042)	(0.050)	
Enrolled Previous Year	0.853	0.835	0.018	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 (OR)	0.460	0.262	0.197	0.000
	(0.039)	(0.034)	(0.038)	

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. They are AND friends if they both name each other as friends.

Table 3.8 presents estimates of Equation (3.6), 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 3.8 shows that if one individual is elected, girls are 9.9 percentage points more likely to be AND friends.

Table 3.8: Baseline Network Links and Election Results

Network Definition	OR (1)	AND (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 Dep Var in Control	0.701	0.286
P-Value of Test 1	0.028	0.003

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 network definition.
Individuals i and j are OR friends if (at least) one names the other as a friend. They are AND friends if they both name each other as friends.
Test 1 is a test of significance of (Elected (OR) + Elected (AND)).

3.3.3 Randomly assigned

As discussed above and presented in Table 3.4, the random assignment of girls to participate in the parliament program resulted in groups that were generally balanced across baseline characteristics.

3.3.4 Program Effects on Networks under different Selection Regimes

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.

3.3.4.1 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.

$$L_{is1}^{linktype} = \delta_0 + \delta_1 T1_s + \delta_2 T2_s + \epsilon_{ijs1} \tag{3.7}$$

Here, $L_{is1}^{linktype}$ indicates the existence of a link between individuals i and j in school s at endline (time 1). The parameters δ_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 3.9 presents results estimates of Equation (3.7). Columns (1) and (4) 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 (4)-(6)).

Table 3.9: ITT Program Effects on Endline Network Formation

Network Definition		OR			AND	
	(1)	(2)	(3)	(4)	(5)	(6)
T1	0.121	0.063	0.013	0.061	-0.015	-0.067
	(0.092)	(0.060)	(0.058)	(0.078)	(0.044)	(0.060)
T2	0.189**	0.131**	0.078	0.180**	0.094	0.053
	(0.076)	(0.056)	(0.050)	(0.076)	(0.059)	(0.065)
Friends at Baseline (OR)		0.158***	0.154***		0.160***	0.157***
		(0.015)	(0.014)		(0.020)	(0.020)
Friends at Baseline (AND)		0.121***	0.120***		0.184***	0.184***
		(0.031)	(0.029)		(0.033)	(0.032)
Constant	0.631***	0.370***	1.105*	0.325***	0.084**	0.923
	(0.071)	(0.067)	(0.586)	(0.065)	(0.039)	(0.703)
Baseline Network Controls	NO	YES	YES	NO	YES	YES
School Controls	NO	NO	YES	NO	NO	YES
Observations	15,578	15,578	15,578	15,578	15,578	15,578
R-squared	0.033	0.137	0.144	0.024	0.161	0.165
Mean Dep Var in Control	0.631	0.631	0.631	0.325	0.325	0.325
P-Value for Test 1	0.296	0.209	0.239	0.056	0.052	0.049

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 under appropriate network definition.
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 the beginning of the school year.
Individuals i and j are OR friends if (at least) one names the other as a friend. They are AND friends if they both name each other as friends.
Test 1 is a test of (T1 - T2 = 0).

3.3.4.2 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:

$$L_{ijs1}^{linktype} = \delta_0 + \delta_1 E_{ijs}^{O} R + \delta_2 E_{ijs}^{AND} + \delta_3 T1_s + \delta_4 T1_s \times E_{ijs}^{OR} + \delta_5 T1_s \times E_{ijs}^{AND}$$
$$+ \delta_6 T1_s + \delta_7 T2_s \times Selected_{ijs}^{OR} + \delta_8 T2_s \times Selected_{ijs}^{AND} + \epsilon_{ijs1} \quad (3.8)$$

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 3.10 are consistent with the baseline network findings in Table 3.8 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 statistically 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 3.10.

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 e ach other, as shown in Table 3.8. 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 participated or were excluded from participation by a pattern consistent with these preexisting networks.

3.3.4.3 Link Formation vs. Retention

The results in Table 3.10 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 (3.9) separately for pairs in each situation.

$$L_{is1}^{linktype} = \delta_0 + \delta_1 T2_s + \delta_2 T2_s \times Selected_{ijs}^{OR} + \delta_3 T2_s \times Selected_{ijs}^{AND} + \epsilon_{ijs1} \tag{3.9}$$

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.

Results are presented in Table 3.11. 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

Table 3.10: Disaggregated ITT Program Effects on Endline Network Formation

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

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 under appropriate network definition.
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 the beginning of the school year.
Individuals i and j are OR friends if (at least) one names the other as a friend. They are AND friends if they both name each other as friends.
Test 1 is a test of the (T1 + T1 × Elected (OR)). Test 2 is a test of the (T1 + T1 × Elected (OR) + T1 × Elected (AND)). Test 3 is a test of (T2 + T2 × Selected (OR)). Test 4 is a test of (T2 + T2 × Selected (OR) + T2 × Selected (AND)).

Table 3.11: Disaggregated ITT Program Effects on Endline 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 × Selected (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 × Selected (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)
R-squared	0.034	0.091	0.106	0.025	0.101	0.115
Mean Dep Var in Control	0.631	0.631	0.631	0.325	0.325	0.325
P-value for Test 1	0.008	0.086	0.680	0.026	0.080	0.588
P-value for Test 2	0.345	0.669	0.781	0.022	0.014	0.083
Panel B: OR Friends but not AND Friends at Baseline						
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 × Selected (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 × Selected (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)
R-squared	0.028	0.083	0.096	0.019	0.063	0.074
Mean Dep Var in Control	0.631	0.631	0.631	0.325	0.325	0.325
P-value for Test 1	0.119	0.080	0.409	0.428	0.318	0.816
P-value for Test 2	0.120	0.184	0.542	0.029	0.009	0.044
Panel C: AND Friends at Baseline						
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 × Selected (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 × Selected (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)
R-squared	0.007	0.045	0.049	0.005	0.079	0.085
Mean Dep Var in Control	0.631	0.631	0.631	0.325	0.325	0.325
P-value for Test 1	0.602	0.591	0.609	0.736	0.909	0.936
P-value for Test 2	0.013	0.011	0.017	0.083	0.184	0.167

Sample restricted to girls in Control and T2 schools. N = 2,544 in Panel A, 4,598 in Panel B, 4,442 in Panel C.
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 under appropriate network definition.
Individuals i and j are OR friends if (at least) one names the other as a friend. They are AND friends if they both name each other as friends.
Test 1 is a test of (T2 + T2 × Selected (OR)). Test 2 is a test of (T2 + T2 × Selected (OR) + T2 × Selected (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 partitioning 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.

3.4 Program Effects on Aspirations and Attitudes

3.4.1 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:

$$y_{is1} = \beta_0 + \beta_1 T1_s + \beta_2 T2_s + \beta_3 y_{is0} + \epsilon_{is1} \tag{3.10}$$

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 3.12 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.

Table 3.12: ITT Program 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
P-value of Test 1	0.438	0.173	0.363	0.345	0.871	0.915	0.834	0.880

Robust standard errors in parentheses, clustered by school. *** p<0.01, ** p<0.05, * p<0.1.
Dependent variable is first principal component of responses to relevant questions on the endline questionnaire.
Test 1 is a test of (T1 - T2 = 0).

3.4.2 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. Aditionally, 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 (3.11).

$$
\begin{aligned}
y_{is1} = {} & \beta_0 + \beta_1 T1_s + \beta_2 T1_s \times Girl_{is} + \beta_3 T1_s \times Girl_{is} \times E_{is} + \beta_4 T2_s \\
& + \beta_5 T2_s \times Girl_{is} + \beta_6 T2_s \times Girl_{is} \times P_{is} + \beta_7 E_{is} + \beta_8 y_{is0} + \epsilon_{is1}
\end{aligned} \tag{3.11}
$$

In Equation (3.11), 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 3.13. 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 $T1_s + T1_s \times Girl_{is}$. 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 3.13 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.

Table 3.13: Disaggregated ITT Program Effects on Endline Attitudes

	Education Index		Career 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 × Selected	-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 Dep Var in C	0.141	0.141	0.041	0.041	0.258	0.258	0.080	0.080
P-value for Test 1	0.138	0.032	0.302	0.302	0.064	0.074	0.878	0.933
P-value for Test 2	0.454	0.156	0.057	0.055	0.381	0.399	0.299	0.275
P-value for Test 3	0.878	0.678	0.402	0.425	0.305	0.298	0.854	0.890
P-value for Test 4	0.556	0.249	0.932	0.916	0.046	0.043	0.287	0.304

Robust standard errors in parentheses, clustered by school. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.
Dependent variable is first principal component of responses to relevant questions on the endline questionnaire.
Test 1 is a test of (T1 + T1 × Girl). Test 2 is a test of (T1 + T1 × Girl + T1 × Elected). Test 3 is a test of (T2 + T2 × Girl). Test 4 is a test of (T2 + T2 × Girl + T2 × Selected).

3.5 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 (Nations, 2015). Social inclusion encompasses a sense of belonging, of integration to the reference group (Shortall, 2008) (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.

APPENDIX A

Proofs of Propositions (Chapter 1)

A.1 Proposition 1.2

Proof. Existence of equilibrium follows directly from Rosen (1965). Given each other player's strategies, each player's uility function is concave in his own strategy g_{is}. Therefore, existence of equilibrium follows from Theorem 1 of Rosen (1965).

I show existence of a strictly positive equilibrium in three steps. First, I show existence of equilibrium in a version of the game in which players' strategy sets are bounded below by $\underline{g} > 0$. Second, I show that, for sufficiently small $\underline{g}$, the lower bound is non-binding. Finally, I demonstrate that the equilibrium of the bounded game is an equilibrium when players are allowed to link zero with other players (that is, when $\underline{g} = 0$).

Step 1: Existence with Strictly Positive Strategy Sets

Define a network-formation game in which individuals maximize utility as defined by Equation (1.11). Different than the game defined in the text, however, they must form strictly positive links with each individual. That is, for each $i, j \neq i$, $g_{ijs} \geq \underline{g}$, where $\underline{g} > 0$ (strictly). Set $\underline{g}$ sufficiently small that each player's strategy set is non-void: $\underline{g} \in (0, \frac{\overline{M}}{(N-1)\underline{c}})$.

As defined by Rosen (1965) and Ui (2008), for each i, $U_{is}(g_{is}, g_{-is})$ is concave for every

g_{-is}. Accordingly, the game is a smooth concave game on a compact strategy set. Thus, by Lemma 1 in Ui (2008) and the notes afterward, a Nash Equilibrium of this game exists.

<u>Step 2</u>: *Lower Bound is Non-Binding for Sufficiently Small* $\underline{g}$

The result in Step 1 applies for any $\underline{g} > 0$ such that strategy sets are non-void. Suppose $\underline{g} \in (0, \frac{\overline{M}}{\underline{c}}(\frac{f_{max}}{f_{min}}\frac{\overline{c}}{\underline{c}})^{\frac{1}{\beta}}) \cap (0, \left(\left(\frac{\underline{M}}{(N-1)\overline{c}}\right)^{\alpha-1}\left(\frac{\overline{M}}{\underline{c}}\right)^{\beta}\frac{f_{max}\overline{c}}{f_{min}\underline{c}}\right)^{\frac{1}{1-\alpha-\beta}})$, where $f_{min} \leq e^{f(X_{is},X_{js})} \leq f_{max} \forall i, j \neq i$.[1] Define λ_{is} as the Lagrange Multiplier for the budget constraint, and μ_{ij} as the Lagrange Multiplier for the lower-bound constraint $g_{ijs} - \underline{g} \geq 0$ for $i, j \neq i$. Therefore, the following Kuhn-Tucker conditions hold for individual i and all $j, k \neq i$:

$$\alpha g_{ijs}^{\alpha-1} g_{jis}^{\beta} e^{f(X_{is},X_{js})} - \lambda_{is} c_{ijs} + \mu_{ijs} = 0 \tag{A.1}$$

$$\alpha g_{iks}^{\alpha-1} g_{kis}^{\beta} e^{f(X_{is},X_{ks})} - \lambda_{is} c_{iks} + \mu_{iks} = 0 \tag{A.2}$$

Suppose the constraint binds for some pair $i, j \neq i$ and thus $g_{ijs} = \underline{g}$. Since utility is increasing in g_{iks} whenever $\underline{g} > 0$, the budget constraint must bind in equilibrium. So, $\sum_{k \neq i} c_{iks} g_{iks} = M_{is}$. Therefore, since $\underline{g} < \frac{\overline{M}}{(N-1)\underline{c}}$, the lower-bound constraint must *not* bind for some $k \neq j, i$. Thus, $g_{iks} > \underline{g}$ and $\mu_{iks} = 0$. Combine Equations (A.1) and (A.2) through λ_{is} as follows:

$$\alpha g_{ijs}^{\alpha-1} g_{jis}^{\beta} \frac{e^{f(X_{is},X_{js})}}{c_{ijs}} + \frac{\mu_{ijs}}{c_{ijs}} = \alpha g_{iks}^{\alpha-1} g_{kis}^{\beta} \frac{e^{f(X_{is},X_{ks})}}{c_{iks}} \tag{A.3}$$

$$g_{ijs}^{\alpha-1} g_{jis}^{\beta} \frac{e^{f(X_{is},X_{js})}}{c_{ijs}} < g_{iks}^{\alpha-1} g_{kis}^{\beta} \frac{e^{f(X_{is},X_{ks})}}{c_{iks}} \tag{A.4}$$

W.l.o.g., choose k such that $g_{iks} \geq g_{ils} \; \forall \; l \neq i$. Due to the budget constraint holding with equality, $g_{iks} \geq \frac{\underline{M}}{(N-1)\overline{c}}$. Since $\alpha - 1 < 0$, $g_{iks}^{\alpha-1} \leq \left(\frac{\underline{M}}{(N-1)\overline{c}}\right)^{\alpha-1}$. Additionally, $g_{kis}^{\beta} \leq \left(\frac{\overline{M}}{\underline{c}}\right)^{\beta}$

[1] f_{min} and f_{max} are well-defined and finite due to compactness of the range of the function f and continuity of the exponential function.

and $\frac{e^{f(X_{is},X_{ks})}}{c_{iks}} \leq \frac{f_{max}}{\underline{c}}$. Substituting on the right-hand side of Equation (A.4) yields

$$g_{ijs}^{\alpha-1} g_{jis}^{\beta} \frac{e^{f(X_{is},X_{js})}}{c_{ijs}} < \left(\frac{\underline{M}}{(N-1)\overline{c}}\right)^{\alpha-1} \left(\frac{\overline{M}}{\underline{c}}\right)^{\beta} \frac{f_{max}}{\underline{c}} \tag{A.5}$$

On the left-hand side of Equation (A.5), $g_{ijs} = \underline{g}$ and $g_{jis} \geq \underline{g}$. Since $\beta > 0$, $g_{ijs}^{\alpha-1} g_{jis}^{\beta} \geq \underline{g}^{\alpha+\beta-1}$. Further, $\frac{e^{f(X_{is},X_{js})}}{c_{ijs}} \geq \frac{f_{min}}{\overline{c}}$. Making these substitutions into Equation (A.5) gives

$$\underline{g}^{\alpha+\beta-1} \frac{f_{min}}{\overline{c}} < \left(\frac{\underline{M}}{(N-1)\overline{c}}\right)^{\alpha-1} \left(\frac{\overline{M}}{\underline{c}}\right)^{\beta} \frac{f_{max}}{\underline{c}} \tag{A.6}$$

Therefore, since $\alpha + \beta - 1 < 0$,

$$\underline{g} > \left(\left(\frac{\underline{M}}{(N-1)\overline{c}}\right)^{\alpha-1} \left(\frac{\overline{M}}{\underline{c}}\right)^{\beta} \frac{f_{max}\overline{c}}{f_{min}\underline{c}}\right)^{\frac{1}{1-\alpha-\beta}} \tag{A.7}$$

This implies a contradiction since we assumed $\underline{g} \in (0, \left(\left(\frac{\underline{M}}{(N-1)\overline{c}}\right)^{\alpha-1} \left(\frac{\overline{M}}{\underline{c}}\right)^{\beta} \frac{f_{max}\overline{c}}{f_{min}\underline{c}}\right)^{\frac{1}{1-\alpha-\beta}})$. Accordingly, for sufficiently small $\underline{g}$, the constraint $g_{ijs} \geq \underline{g}$ does not hold for any pair $i, j \neq i$. Therefore, with this restriction, there exists an equilibrium in which $g_{ijs} > \underline{g}\ \forall\ i, j \neq i$ (strictly).

<u>Step 3</u>: Equilibrium of the bounded game is still an equilibrium when players are allowed to choose links of 0.

Step 2 above shows that, for arbitrarily small $\underline{g}$, an equilibrium exists in which the lower-bound condition is non-binding for every pair $i, j \neq i$. Therefore, for all i, any deviation in which $g'_{ijs} > 0\ \forall\ i, j \neq i$ cannot lead to higher utility to i than this equilibrium allocation. I now show that this fact remains true if we allow players to choose $g'_{ijs} = 0$.

Assume players play the strategies played in the equilibrium described in Step 2. For every $i, j \neq i$, define this strategy as g_{ijs}. At this point, $\mu_{ijs} = 0\ \forall\ i, j \neq i$, and Equation

(A.1) becomes

$$\alpha g_{ijs}^{\alpha-1} g_{jis}^{\beta} e^{f(X_{is},X_{js})} = \lambda_{is} c_{ijs} \tag{A.8}$$

From this, we see that

$$U_i(g_{is}, g_{-is}) = \sum_{j \neq i} g_{ijs}^{\alpha} g_{jis}^{\beta} e^{f(X_{is},X_{js})} = \frac{1}{\alpha} \sum_{j \neq i} \lambda_{is} c_{ijs} g_{ijs} \tag{A.9}$$

$$= \lambda_{is} \frac{M_{is}}{\alpha} \tag{A.10}$$

Suppose that this is not an equilibrium of the game in which $\underline{g} = 0$. Therefore, for some i, there exists an alternative strategy g'_{is} in which $g'_{ijs} = 0$ for some $j \neq i$ where $U_i(g'_{is}, g_{-is}) > U_i(g_{is}, g_{-is})$.

The utility from links where $g'_{ijs} > 0$ is bounded above by the utility derived from solving the First Order Conditions in Equation (A.8), restricted to positive links. Define g''_{ijs} as the hypothetical set of links in which these FOCs hold whenever $g'_{ijs} > 0$. Note that $g''_{ijs} > \iff g'_{ijs} > 0$, and λ''_{is} as the Lagrange Multiplier corresponding to this constrained utility-maximizing strategy. So,

$$\alpha (g''_{ijs})^{\alpha-1} g_{jis}^{\beta} e^{f(X_{is},X_{js})} = \lambda''_{is} c_{ijs} \; \forall \; j \neq i \mid g'_{ijs} > 0 \tag{A.11}$$

Therefore,

$$\frac{g''_{ijs}}{g_{ijs}} = \left(\frac{\lambda''_{is}}{\lambda_{is}} \right)^{\frac{1}{\alpha-1}} \tag{A.12}$$

From this, we see that $\frac{g''_{ijs}}{g_{ijs}}$ is constant across all j for whom $g''_{ijs} > 0$. Clearly, $g''_{ijs} > g_{ijs}$

whenever $g''_{ijs} > 0$. Further, $0 < \alpha < 1 \Rightarrow \lambda''_{is} < \lambda_{is}$. From this, we see that

$$
\begin{aligned}
U_{ijs}(g'_{is}, g_{-is}) &= \sum_{j \neq i} (g'_{ijs})^{\alpha} g^{\beta}_{jis} e^{f(X_{is}, X_{js})} && \text{(A.13)} \\
&= \sum_{j \neq i} 1\{g'_{ijs} > 0\} (g'_{ijs})^{\alpha} g^{\beta}_{jis} e^{f(X_{is}, X_{js})} && \text{(A.14)} \\
&\leq \frac{1}{\alpha} \sum_{j \neq i} 1\{g'_{ijs} > 0\} \lambda''_{is} g''_{ijs} c_{ijs} && \text{(A.15)} \\
&= \lambda''_{is} \frac{M_{is}}{\alpha} && \text{(A.16)} \\
&< \lambda_{is} \frac{M_{is}}{\alpha} = U_i(g_{is}, g_{-is}) && \text{(A.17)}
\end{aligned}
$$

This shows that *any* deviation in which $g_{ijs} = 0$ for some j makes agent i strictly worse off. Therefore, the strictly positive equilibrium is also an equilibrium of the game when $\underline{g} = 0$.

□

A.2 Proposition 1.3

Proof. Suppose there are two equilibria (g, λ) and (g', λ'), where $g = (g_{12s}, g_{13s}, ..., g_{NN-1s})$ and $\lambda = (\lambda_{1s}, ..., \lambda_{Ns})$. Equations (1.12) and (1.13), the First Order necessary conditions for strictly positive equilibrium, imply

$$
(\alpha - 1)(\log g_{ijs} - \log g'_{ijs}) + \beta(\log g_{jis} - \log g'_{jis}) - (\log \lambda_{is} - \log \lambda'_{is}) = 0 \; \forall \, i, j \neq i \quad \text{(A.18)}
$$

$$
\sum_{j \neq i} c_{ijs}(g_{ijs} - g'_{ijs}) = 0 \; \forall \, i \quad \text{(A.19)}
$$

Define $\tilde{\beta} = \frac{\beta}{1-\alpha}$ and $\tilde{\lambda}_{is} = \frac{\log \lambda_{is}}{1-\alpha}$. After substitution and rearrangement, Equation (A.18) becomes

$$
(\log g_{ijs} - \log g'_{ijs}) = \tilde{\beta}(\log g_{jis} - \log g'_{jis}) - (\tilde{\lambda}_{is} - \tilde{\lambda}'_{is}) \; \forall \, i, j \neq i \quad \text{(A.20)}
$$

By symmetry,

$$(\log g_{jis} - \log g'_{jis}) = \tilde{\beta}(\log g_{ijs} - \log g'_{ijs}) - (\tilde{\lambda}_{js} - \tilde{\lambda}'_{js}) \; \forall \; i, j \neq i \tag{A.21}$$

Substitute Equation (A.21) into Equation (A.20) and rearrange, yielding

$$(\log g_{ijs} - \log g'_{ijs}) = -\frac{1}{1-\tilde{\beta}^2}\left(\tilde{\beta}(\tilde{\lambda}_{js} - \tilde{\lambda}'_{js}) + (\tilde{\lambda}_{is} - \tilde{\lambda}'_{is})\right) \; \forall \; i, j \neq i \tag{A.22}$$

Since the log function is continuously differntiable for all positive values, the Mean Value Theorem $\Rightarrow \exists \; g^*_{ijs} \in [g_{ijs}, g'_{ijs}]$, where $\log g_{ijs} - \log g'_{ijs} = \frac{1}{g^*_{ijs}}(g_{ijs} - g'_{ijs})$ and $g^*_{ijs} > 0$. Make this substitution and multiply by $-(1-\tilde{\beta}^2)g^*_{ijs}c_{ijs}$:

$$-(1-\tilde{\beta}^2)c_{ijs}(g_{ijs} - g'_{ijs}) = c_{ijs}g^*_{ijs}\left(\tilde{\beta}(\tilde{\lambda}_{js} - \tilde{\lambda}'_{js}) + (\tilde{\lambda}_{is} - \tilde{\lambda}'_{is})\right) \; \forall \; i, j \neq i \tag{A.23}$$

Next, sum across $j \neq i$, substitute and rearrange:

$$-(1-\tilde{\beta}^2)\sum_{j\neq i} c_{ijs}(g_{ijs} - g'_{ijs}) = \sum_{j\neq i} c_{ijs}g^*_{ijs}\left(\tilde{\beta}(\tilde{\lambda}_{js} - \tilde{\lambda}'_{js}) + (\tilde{\lambda}_{is} - \tilde{\lambda}'_{is})\right) \; \forall \; i \tag{A.24}$$

$$0 = \left(\sum_{j\neq i} c_{ijs}g^*_{ijs}\right)(\tilde{\lambda}_{is} - \tilde{\lambda}'_{is}) + \tilde{\beta}\sum_{j\neq i} c_{ijs}g^*_{ijs}(\tilde{\lambda}_{js} - \tilde{\lambda}'_{js}) \; \forall \; i \tag{A.25}$$

This defines a linear system of N equations and N unknowns, as defined by $\mathbf{Ab} = 0$ in Equation (A.26):

$$\begin{bmatrix} (\sum_{j\neq 1} c_{1js}g^*_{1js}) & \tilde{\beta}c_{12s}g^*_{12s} & ... & \tilde{\beta}c_{1Ns}g^*_{1Ns} \\ \tilde{\beta}c_{21s}g^*_{21s} & (\sum_{j\neq 2} c_{2js}g^*_{2js}) & ... & \tilde{\beta}c_{2Ns}g^*_{2Ns} \\ . & . & . & . \\ \tilde{\beta}c_{N1s}g^*_{N1s} & . & ... & (\sum_{j\neq 1} c_{Njs}g^*_{Njs}) \end{bmatrix} \begin{bmatrix} \tilde{\lambda}_{1s} - \tilde{\lambda}'_{1s} \\ \tilde{\lambda}_{2s} - \tilde{\lambda}'_{2s} \\ . \\ . \\ \tilde{\lambda}_{Ns} - \tilde{\lambda}'_{Ns} \end{bmatrix} = 0 \tag{A.26}$$

Clearly, $\mathbf{A}$ being invertible will guarantee $\tilde{\lambda}_{is} - \tilde{\lambda}'_{is} = 0 \; \forall \; i$.

Suppose $\mathbf{A}$ is not invertible. Therefore, 0 is an eigenvalue of $\mathbf{A}$ with an associated eigenvector $\mathbf{v}$. Let v_m be the largest element of $\mathbf{v}$ and, w.l.o.g., $v_m > 0$. So, $v_m \geq v_j \geq -v_m \; \forall j \neq m$. Now,

$$v_m(\sum_{j \neq m} c_{mjs} g^*_{mjs}) + \tilde{\beta} \sum_{j \neq m} v_j c_{mjs} g^*_{mjs} \geq v_m(\sum_{j \neq m} c_{mjs} g^*_{mjs}) - \tilde{\beta} v_m \sum_{j \neq m} c_{mjs} g^*_{mjs} \tag{A.27}$$

$$> v_m(\sum_{j \neq m} c_{mjs} g^*_{mjs})(1 - \tilde{\beta}) > 0 \tag{A.28}$$

This contradicts that 0 is an eigenvalue. Therefore, $\mathbf{A}$ is invertible, and $\tilde{\lambda}_{is} = tilde\lambda'_{is} \forall i$.

Finally, from Equation (A.22), we see that $\tilde{\lambda}_{is} - \tilde{\lambda}'_{is} = 0 \; \forall \; i, j \neq i \; \Rightarrow (\log g_{ijs} - \log g'_{ijs}) = 0 \; \forall \; i, \; j \neq i$. Therefore, $(g, \lambda) = (g', \lambda')$ and the equilibrium is unique.

□

A.3 Proposition 1.4

Proof. Let $\mathbf{g_{ks}} = (g_{k1s}, ..., g_{kNs})$ be agent i's strategy vector, and $\mathbf{g_{-is}}$ be the strategy vectors of the other $N_s - 1$ players. The definition of a potential game requires that, for every $k, \mathbf{g_{ks}}, \mathbf{g_{-ks}}$,

$$P(\mathbf{g_{ks}}, \mathbf{g_{-ks}}, \mathbf{X_s}) - P(\mathbf{g'_{ks}}, \mathbf{g_{-ks}}, \mathbf{X_s}) = U_{is}(\mathbf{g_{ks}}, \mathbf{g_{-ks}}, \mathbf{X_s}) - U_{is}(\mathbf{g'_{ks}}, \mathbf{g_{-ks}}, \mathbf{X_s})$$

Simple substitution and the assumption f $(f(X_{is}, X_{ks}) = f(X_{ks}, X_{is})$ shows this to be the case. Define $u_{ijs}(\mathbf{g_{ks}}, \mathbf{g_{-ks}}) = (g_{ijs} g_{jis})^{\alpha} e^{f(X_{is}, X_{js})}$ and $u_{ijs}(\mathbf{g'_{ks}}, \mathbf{g_{-ks}})$ similarly. Note that

$u_{ijs}(\mathbf{g_{ks}}, \mathbf{g_{-ks}}) = u_{ijs}(\mathbf{g'_{ks}}, \mathbf{g_{-ks}})$ whenever $i, j \neq k$. So,

$$\begin{aligned}
P(\mathbf{g_{ks}}, \mathbf{g_{-ks}}, \mathbf{X_s}) - P(\mathbf{g'_{ks}}, \mathbf{g_{-ks}}, \mathbf{X_s}) &= \frac{1}{2}\sum_{i=1}^{N_s}\sum_{j\neq i}\left(u_{ijs}(\mathbf{g_{ks}}, \mathbf{g_{-ks}}) - u_{ijs}(\mathbf{g'_{ks}}, \mathbf{g_{-ks}})\right) \\
&= \frac{1}{2}\sum_{j\neq k}\left(u_{kjs}(\mathbf{g_{ks}}, \mathbf{g_{-ks}}) - u_{kjs}(\mathbf{g'_{ks}}, \mathbf{g_{-ks}})\right) \\
&+ \frac{1}{2}\sum_{i\neq k}\left(u_{iks}(\mathbf{g_{ks}}, \mathbf{g_{-ks}}) - u_{iks}(\mathbf{g'_{ks}}, \mathbf{g_{-ks}})\right) \\
&= \sum_{j\neq k}\left(u_{jks}(\mathbf{g_{ks}}, \mathbf{g_{-ks}}) - u_{jks}(\mathbf{g'_{ks}}, \mathbf{g_{-ks}})\right) \\
&= \sum_{j\neq k}(g_{kjs}g_{jks})^{\alpha}e^{f(X_{ks},X_{js})} - \sum_{j\neq k}(g'_{kjs}g_{jks})^{\alpha}e^{f(X_{ks},X_{js})} \\
&= U_{is}(\mathbf{g_{ks}}, \mathbf{g_{-ks}}, \mathbf{X_s}) - U_{is}(\mathbf{g'_{ks}}, \mathbf{g_{-ks}}, \mathbf{X_s})
\end{aligned}$$

□

A.4 Proposition 1.5

Proof. First, it is clear that the potential function $P(\mathbf{G_s}, \mathbf{X_s})$ is a continuous function in $\mathbf{G_s}$ on a compact set. Therefore, there exists some $\mathbf{G_s^*}$ that maximizes the potential function.

Monderer and Shapley (1996) showed that the set of strategy profiles that maximize P is a subset of the set of equilibrium profiles. Since the game as a potential game is a special case of the broader game, Proposition 1.2 provides existence results, and Proposition 1.3 provides that the strictly positive equilibrium is unique. Therefore, by showing that any other equilibrium is *not* a potential function maximizer, by necessity the strictly positive one must be.

I prove this by demonstrating that any equilibrium strategy profile that is *not* the strictly positive equilibrium cannot maximize P.

Take any equilibrium network $\mathbf{G_s}$ *other than* the strictly positive one. Therefore, $g_{ijs} =$

$g_{jis} = 0$ for some $i, j \neq i$. For all $d \in [0, \frac{M_{is}}{c_{ijs}}] \cap [0, \frac{M_{js}}{c_{jss}}]$, define a deviation profile as follows:

$$\begin{aligned}
g'_{ijs} &= g'_{jis} = d \\
g'_{iks} &= \frac{M_{is} - dc_{ijs}}{M_{is}} g_{iks} \;\forall\; k \neq i, j \\
g'_{jks} &= \frac{M_{js} - dc_{jis}}{M_{js}} g_{jks} \;\forall\; k \neq i, j \\
g'_{kls} &= g_{kls} \;\forall\; k \neq i, j, \forall l
\end{aligned}$$

That is, g_{iks} and g_{jks} adjust proportionally. Note that $g_{iks} = 0 \Rightarrow g'_{iks} = 0$. Such a deviation is feasible (within the constraint set) for all d as restricted above.

Now, define a function $F(d)$ as follows:

$$\begin{aligned}
F(d) &= P(\mathbf{G_s}', \mathbf{X_s}) - P(\mathbf{G_s}, \mathbf{X_s}) \\
&= d^{2\alpha} e^{f(X_{is}, X_{js})} + ((\frac{M_{is} - dc_{ijs}}{M_{is}})^{\alpha} - 1) U_{is}(\mathbf{G_s}, \mathbf{X_s}) + ((\frac{M_{js} - dc_{jis}}{M_{js}})^{\alpha} - 1) U_{js}(\mathbf{G_s}, \mathbf{X_s})
\end{aligned} \tag{A.29}$$

For notational convenience, let $U_{is} = U_{is}(\mathbf{G_s}, \mathbf{X_s})$ and $U_{js} = U_{js}(\mathbf{G_s}, \mathbf{X_s})$. The function $F(d)$ is continuous within its domain and differentiable for all $d > 0$. Further,

$$F'(d) = 2\alpha d^{2\alpha-1} e^{f(X_{is}, X_{js})} - \frac{\alpha c_{ijs} U_{is}}{M_{is}} (\frac{M_{is} - dc_{ijs}}{M_{is}})^{\alpha-1} - \frac{\alpha c_{jis} U_{js}}{M_{js}} (\frac{M_{js} - dc_{jis}}{M_{js}})^{\alpha-1} \tag{A.30}$$

$$\geq 2\alpha d^{2\alpha-1} e^{f(X_{is}, X_{js})} - \frac{\alpha \bar{c}}{\underline{M}} \left(U_{is} (\frac{M_{is}}{M_{is} - dc_{ijs}})^{1-\alpha} + U_{is} (\frac{M_{js}}{M_{js} - dc_{jis}})^{1-\alpha} \right) \tag{A.31}$$

Next, I show that there exists d such that $F'(d)$ is strictly positive. There are two distinct cases: (1) $U_{is} = U_{js} = 0$, and (2) $U_{is} + U_{js} > 0$.

Case 1: $U_{is} = U_{js} = 0$

In this case, Equation (A.31) becomes

$$F'(d) = 2\alpha d^{2\alpha-1} e^{f(X_{is}, X_{js})} \tag{A.32}$$

which is strictly positive for all $d > 0$. Therefore, by the Mean Value Theorem, there exists some $d^* \in (0, d)$ such that $F(d) - F(0) = F'(d^*)(d - 0)$. So, $F(d) > 0$.

Case 2: $U_{is} + U_{js} > 0$

When $U_{is}+U_{js} > 0$, further restrict d s.t. $d \in (0, \frac{M_{is}}{2c_{ijs}}) \cap (0, \frac{M_{is}}{2c_{ijs}}) \cap (0, \left(\frac{\overline{c}(U_{is}+U_{js})}{\underline{M}e^{f(X_{is},X_{js})}}\right)^{\frac{1}{1-2\alpha}})$.[2] This set is non-void since the supremum of each interval is a strictly positive number. Within this restricted set,

$$\begin{aligned} F'(d) &> 2\alpha \left(\frac{\overline{c}(U_{is}+U_{js})}{\underline{M}e^{f(X_{is},X_{js})}}\right) e^{f(X_{is},X_{js})} - \frac{\alpha\overline{c}}{\underline{M}} 2^{1-\alpha}(U_{is}+U_{js}) \\ &= \frac{\alpha\overline{c}(U_{is}+U_{js})}{\underline{M}}(2 - 2^{1-\alpha}) > 0 \end{aligned} \tag{A.33}$$

Therefore, for any $d \in (0, \frac{M_{is}}{2c_{ijs}}) \cap (0, \frac{M_{is}}{2c_{ijs}}) \cap (0, \left(\frac{\overline{c}(U_{is}+U_{js})}{\underline{M}e^{f(X_{is},X_{js})}}\right)^{\frac{1}{1-2\alpha}})$, by the Mean Value Theorem, there exists some $d^* \in (0, d)$ such that $F(d) - F(0) = F'(d^*)(d - 0)$. So, $F(d) > 0$.

Bringing it all together, for *any* equilibrium $\mathbf{G_s}$ where $g_{ijs} = g_{jis} = 0$ for some $i, j \neq i$, there exists some feasible deviation $\mathbf{G'_s}$ in which $F(d) > 0$ and thus $P(\mathbf{G'_s}, \mathbf{X_s}) > P(\mathbf{G_s}, \mathbf{X_s})$. Therefore, each such $\mathbf{G_s}$ is not a maximizer of the potential function. Accordingly, the unique strictly positive strategy profile, for which $g_{ijs} > 0 \;\forall\, i, j \neq i$, maximizes the potential function.

□

A.5 Proposition 1.6

Proof. This result is a typical panel IV result, allowing for arbitrary correlation of variables within clusters. Let S be the number of schools (potential networks) observed and N be the number of actors per school. Starting with Equation (1.20) and the instrument set z_{ijs}, we

[2] In this case where the game is a potential game, $\beta = \alpha$. The assumption 1.2 therefore implies that $1 - 2\alpha > 0$.

see that

$$
\begin{aligned}
z'_{ijs}\dot{g}^i_{ijs} &= z'_{ijs}\dot{g}^i_{jis}\tilde{\beta} + z'_{ijs}X_{is} \odot \dot{X}^i_{js}\delta_1 + z'_{ijs}\dot{X}^i_{js}\gamma_3 - z'_{ijs}\dot{c}^i_{ijs} \\
&= z'_{ijs}b_{ijs}\theta - z'_{ijs}\dot{c}^i_{ijs} \qquad \text{(A.34)}
\end{aligned}
$$

where $\theta = (\tilde{\beta}, \delta'_1, \gamma'_3)'$. Next, sum across all schools and pairs of students:

$$
\begin{aligned}
\frac{1}{SN(N-1)} \sum_{s=1}^{S} \sum_{i=1}^{N} \sum_{j \neq i} b'_{ijs} z_{ijs} z'_{ijs} \dot{g}^i_{ijs} &= \sum_{s=1}^{S} \sum_{i=1}^{N} \sum_{j \neq i} \frac{1}{SN(N-1)} b'_{ijs} z_{ijs} z'_{ijs} b_{ijs} \theta \\
&\quad - \frac{1}{Sn(n-1)} \sum_{s=1}^{S} \sum_{i=1}^{N} \sum_{j \neq i} b'_{ijs} z_{ijs} z'_{ijs} \dot{c}^i_{ijs} \qquad \text{(A.35)}
\end{aligned}
$$

Since $\dot{c}^i_{ijs}$ is a linear combination of terms that are assumed to be independent across s, $\dot{c}^i_{ijs}$ is also independent across s. Further, all terms are bounded and thus have finite variance. Let w_{ijs} be an element of one of the matrices in Equation (A.35). For any such variable, $\mathbb{C}[w_{ijs}, w_{klt}] = 0$ whenever $s \neq t$. Further, since each w_{ijs} is identically distributed within a school, $\mathbb{V}[w_{ijs}] = \mathbb{V}[w_{lks}] \forall i, j, k, l$. Further,

$$
\begin{aligned}
\mathbb{V}[\bar{w}_{ijs}] &= \mathbb{V}[\frac{1}{SN(N-1)} \sum_{s=1}^{S} \sum_{i=1}^{N} \sum_{j \neq i} w_{ijs}] \\
&= \frac{1}{S^2N^2(N-1)^2} \left(SN(N-1)\mathbb{V}[w_{ijs}] + S(\sum_{i=1}^{N} \sum_{j \neq i})(\sum_{k=1}^{N} \sum_{l \neq j})\mathbb{C}[w_{ijs}, w_{kls}] \right) \\
&\leq \frac{1}{S} \left(\frac{1}{N(N-1)} + 1 \right) \mathbb{V}[w_{ijs}] \qquad \text{(A.36)}
\end{aligned}
$$

where the final line applies the Cauchy Schwarz Inequality ($\mathbb{C}[w_{ijs}, w_{kls}] \leq \mathbb{V}[w_{ijs}]$). Therefore, $\lim_{S\to\infty} \mathbb{V}[\bar{w}_{ijs}] = 0$ and by Chebyshev's Inequality,

$$
\operatorname*{plim}_{S\to\infty} \bar{w}_{ijs} = \mathbb{E}[w_{ijs}] \qquad \text{(A.37)}
$$

I here note that all terms in Equation (A.35) are sample averages. Therefore, we can apply Equation (A.37) to each element. Thus, $\frac{1}{SN(N-1)}\sum_{s=1}^{S}\sum_{i=1}^{N}\sum_{j\neq i} b'_{ijs}z_{ijs}z'_{ijs}\dot{g}^i_{ijs} \rightarrow_p$ $\mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}\dot{g}^i_{ijs}]$, etc. Now, replacing the terms in Equation (A.35) with probability limits,

$$\mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}\dot{g}^i_{ijs}] = \mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}b_{ijs}]\theta - \mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}\dot{c}^i_{ijs}] \tag{A.38}$$

Assumption 1.4 implies that the final term in Equation (A.38) is zero, while the rank condition of Proposition 1.6 implies invertibility of $\mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}b_{ijs}]$. Therefore,

$$\theta = (\mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}b_{ijs}])^{-1}\mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}\dot{g}^i_{ijs}] \tag{A.39}$$

and thus the parameters $\tilde{\beta}$, δ_1, and γ_3 are identified.

□

A.6 Proposition 1.7

Proof. This proof is very similar to Proposition 1.6 but relies upon the additional exogeneity conditions in Assumption 1.6. Starting with Equation (1.27) and the instrument set z_{ijs}, we see

$$\begin{aligned} z'_{ijs}\dot{g}^i_{ijs} &= z'_{ijs}\dot{g}^i_{jis}\tilde{\beta} + z'_{ijs}X_{is}\dot{X}^i_{js}\delta_1 + z'_{ijs}X_{is}\dot{a}^i_{js}\delta_2 + z'_{ijs}a_{is}\dot{X}^i_{js}\delta_3 + z'_{ijs}a_{is}\dot{a}^i_{js}\delta_4 \\ &\quad + z'_{ijs}\dot{X}^j_{is}\gamma_3 + z'_{ijs}a_{is}\dot{a}^i_{js}\gamma_4 - z'_{ijs}\dot{c}^i_{ijs} \end{aligned} \tag{A.40}$$

Rearrangement of terms shows that

$$z'_{ijs}\dot{g}^i_{ijs} = z'_{ijs}b_{ijs}\theta + z'_{ijs}X_{is}\dot{a}^i_{js}\delta_2 + z'_{ijs}a_{is}\dot{X}^i_{js}\delta_3 + z'_{ijs}a_{is}\dot{a}^i_{js}\delta_4 + z'_{ijs}a_{is}\dot{a}^i_{js}\gamma_4 - z'_{ijs}\dot{c}^i_{ijs} \tag{A.41}$$

where $\theta = (\tilde{\beta}, \delta_1', \gamma_3)'$. Next, sum across all schools and all pairs of students. So,

$$\begin{aligned}\frac{1}{SN(N-1)}\sum_{s=1}^{S}\sum_{i=1}^{N}\sum_{j\neq i} b_{ijs}' z_{ijs} z_{ijs}' \dot{g}_{ijs}^i = \frac{1}{SN(N-1)}\sum_{s=1}^{S}\sum_{i=1}^{N}\sum_{j\neq i} (&b_{ijs}' z_{ijs} z_{ijs}' b_{ijs}\theta \\ &+ b_{ijs}' z_{ijs} z_{ijs}' X_{is}\dot{a}_{js}^i \delta_2 + b_{ijs}' z_{ijs} z_{ijs}' a_{is}\dot{X}_{js}^i \delta_3 \\ &+ b_{ijs}' z_{ijs} z_{ijs}' a_{is}\dot{a}_{js}^i \delta_4 + b_{ijs}' z_{ijs} z_{ijs}' a_{is}\dot{a}_{js}^i \gamma_4 \\ &- b_{ijs}' z_{ijs} z_{ijs}' \dot{c}_{ijs}^i)\end{aligned} \tag{A.42}$$

Since $\dot{c}_{ijs}^i$ is a linear combination of terms that are assumed to be independent across s, $\dot{c}_{ijs}^i$ is also independent across s. Further, all terms are bounded and thus have finite variance. Let w_{ijs} be an element of one of the matrices in Equation (A.42). For any such variable, $\mathbb{C}[w_{ijs}, w_{klt}] = 0$ whenever $s \neq t$. Further,

$$\begin{aligned}\mathbb{V}[\bar{w}_{ijs}] &= \mathbb{V}[\frac{1}{SN(N-1)}\sum_{s=1}^{S}\sum_{i=1}^{N}\sum_{j\neq i} w_{ijs}] \\ &= \frac{1}{S^2N^2(N-1)^2}\left(SN(N-1)\mathbb{V}[w_{ijs}] + S(\sum_{i=1}^{N}\sum_{j\neq i})(\sum_{k=1}^{N}\sum_{l\neq j})\mathbb{C}[w_{ijs}, w_{kls}]\right) \\ &\leq \frac{1}{S}\left(\frac{1}{N(N-1)} + 1\right)\mathbb{V}[w_{ijs}]\end{aligned} \tag{A.43}$$

where the final line applies the Cauchy-Schwarz Inequality ($\mathbb{C}[w_{ijs}, w_{kls}] \leq \mathbb{V}[w_{ijs}]$). Therefore, $\lim_{S\to\infty}\mathbb{V}[\bar{w}_{ijs}] = 0$ and by Chebyshev's Inequality,

$$\operatorname*{plim}_{S\to\infty} \bar{w}_{ijs} = \mathbb{E}[w_{ijs}] \tag{A.44}$$

I here note that all terms in Equation (A.42) are sample averages. Therefore, we can apply Equation (A.44) to each element. Thus, $\frac{1}{SN(N-1)}\sum_{s=1}^{S}\sum_{i=1}^{N}\sum_{j\neq i} b_{ijs}' z_{ijs} z_{ijs}' \dot{g}_{ijs}^i \to_p \mathbb{E}[b_{ijs}' z_{ijs} z_{ijs}' \dot{g}_{ijs}^i]$, etc. Now, replacing the terms in Equation (A.42) with probability limits,

$$\mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}\dot{g}^i_{ijs}] = \mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}b_{ijs}]\theta + \mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}X_{is}\dot{a}^i_{js}]\delta_2 + \mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}a_{is}\dot{X}^i_{js}]\delta_3$$
$$+ \mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}a_{is}\dot{a}^i_{js}]\delta_4 + \mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}\dot{a}^i_{js}]\gamma_4 - \mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}\dot{c}^i_{ijs}] \quad \text{(A.45)}$$

The first part of Assumption 1.6 implies that the final term in Equation (A.45) is zero. Note that z_{ijs} and b_{ijs} are simply functions of x_{ks}. Therefore, application of L.I.E. implies that $\mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}X_{is}\dot{a}^i_{js}] = \mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}X_{is}\mathbb{E}[\dot{a}^i_{js}|b_{ijs}, z_{ijs}]] = 0$. By similar argument, $\mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}a_{is}\dot{X}^i_{js}] = 0$ and $\mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}\dot{a}^i_{js}] = 0$. Further, the third part of Assumption 1.6 implies

$$\begin{aligned}
\mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}a_{is}\dot{a}^i_{js}] &= \mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}a_{is}\dot{a}^i_{js}] \\
&= \mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}\mathbb{E}[\dot{a}^i_{js}a_{is}|x_{ks}]] \\
&= \mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}\mathbb{E}[\dot{a}^i_{js}\mathbb{E}[a_{is}|x_{ks}, a_{ls}]|x_{ks}]] \\
&= 0 \qquad \text{(A.46)}
\end{aligned}$$

where we condition on all k and $l \neq i$. Substituting these results into Equation (A.45) shows that

$$\mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}\dot{g}^i_{ijs}] = \mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}b_{ijs}]\theta \quad \text{(A.47)}$$

The rank condition guarantees the existence of $(\mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}b_{ijs}])^{-1}$ and thus

$$\theta = (\mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}b_{ijs}])^{-1}\mathbb{E}[b'_{ijs}z_{ijs}z'_{ijs}\dot{g}^i_{ijs}] \quad \text{(A.48)}$$

So, $\theta = (\tilde{\beta}, delta'_1, \gamma'_3)$ is identified. □

A.7 Proposition 1.8

The first rank condition, together with Assumption 1.6 and Proposition 1.7, imply that $\tilde{\beta}$ is identified. I prove the rest of the proposition in three steps: (1) Scale identification of δ_2 and γ_2, (2) Scale identification of δ_3, and (3) Scale identification of δ_4.

Step 1: Scale identification of δ_2 and γ_2

Proof. To begin, multiply Equation (1.27) by $z'_{ijs}a_{js}$ and sum across $SN(N-1)$ observations. So,

$$\frac{1}{SN(N-1)}\sum_{s=1}^{S}\sum_{i=1}^{N}\sum_{j\neq i} z'_{ijs}a_{js}\dot{g}^{i}_{ijs} = \frac{1}{SN(N-1)}\sum_{s=1}^{S}\sum_{i=1}^{N}\sum_{j\neq i}(z'_{ijs}a_{js}\dot{g}^{i}_{jis}\tilde{\beta} + z'_{ijs}a_{js}X_{is}\dot{X}^{i}_{js}\delta_1$$
$$+ z'_{ijs}a_{js}X_{is}\dot{a}^{i}_{js}\delta_2 + z'_{ijs}a_{js}a_{is}\dot{X}^{i}_{js}\delta_3 + z'_{ijs}a_{js}a_{is}\dot{a}^{i}_{js}\delta_4$$
$$+ z'_{ijs}a_{js}\dot{X}^{j}_{is}\gamma_3 + z'_{ijs}a_{js}a_{is}\dot{a}^{i}_{js}\gamma_4 - z'_{ijs}a_{js}\dot{c}^{i}_{ijs}) \tag{A.49}$$

Since $\dot{c}^{i}_{ijs}$ is a linear combination of terms that are assumed to be independent across s, $\dot{c}^{i}_{ijs}$ is also independent across s. Further, all terms are bounded and thus have finite variance. Let w_{ijs} be an element of one of the matrices in Equation (A.49). For any such variable, $\mathbb{C}[w_{ijs}, w_{klt}] = 0$ whenever $s \neq t$. Further,

$$\begin{aligned}
\mathbb{V}[\bar{w}_{ijs}] &= \mathbb{V}[\frac{1}{SN(N-1)}\sum_{s=1}^{S}\sum_{i=1}^{N}\sum_{j\neq i} w_{ijs}] \\
&= \frac{1}{S^2N^2(N-1)^2}\left(SN(N-1)\mathbb{V}[w_{ijs}] + S(\sum_{i=1}^{N}\sum_{j\neq i})(\sum_{k=1}^{N}\sum_{l\neq j})\mathbb{C}[w_{ijs}, w_{kls}]\right) \\
&\leq \frac{1}{S}\left(\frac{1}{N(N-1)} + 1\right)\mathbb{V}[w_{ijs}]
\end{aligned} \tag{A.50}$$

where the final line applies the Cauchy-Schwarz Inequality ($\mathbb{C}[w_{ijs}, w_{kls}] \leq \mathbb{V}[w_{ijs}]$). Therefore, $\lim_{S\to\infty} \mathbb{V}[\bar{w}_{ijs}] = 0$ and by Chebyshev's Inequality,

$$\operatorname*{plim}_{S\to\infty} \bar{w}_{ijs} = \mathbb{E}[w_{ijs}] \tag{A.51}$$

I here note that all terms in Equation (A.42) are sample averages. Therefore, we can apply Equation (A.51)to each element. Thus, $\frac{1}{SN(N-1)} \sum_{s=1}^{S} \sum_{i=1}^{N} \sum_{j\neq i} z'_{ijs} a_{js} \dot{g}^i_{ijs} \to_p \mathbb{E}[z'_{ijs} a_{js} \dot{g}^i_{ijs}]$, etc. Now, replacing the terms in Equation (A.42) with probability limits,

$$\begin{aligned}\mathbb{E}[z'_{ijs} a_{js} \dot{g}^i_{ijs}] = \mathbb{E}[z'_{ijs} a_{js} \dot{g}^i_{jis}]\tilde{\beta} + \mathbb{E}[z'_{ijs} a_{js} X_{is} \dot{X}^i_{js}]\delta_1 + \mathbb{E}[z'_{ijs} a_{js} X_{is} \dot{a}^i_{js}]\delta_2 + \mathbb{E}[z'_{ijs} a_{js} a_{is} \dot{X}^i_{js}]\delta_3 \\ + \mathbb{E}[z'_{ijs} a_{js} a_{is} \dot{a}^i_{js}]\delta_4 + \mathbb{E}[z'_{ijs} a_{js} \dot{X}^j_{is}]\gamma_3 + \mathbb{E}[z'_{ijs} a_{js} a_{is} \dot{a}^i_{js} \gamma_4] - \mathbb{E}[z'_{ijs} a_{js} \dot{c}^i_{ijs}]\end{aligned} \tag{A.52}$$

Assumption 1.6 implies that $\mathbb{E}[z'_{ijs} a_{js} \dot{c}^i_{ijs}] = 0$.

Assumption 1.6 and L.I.E. imply that $\mathbb{E}[z'_{ijs} a_{js} X_{is} \dot{X}^i_{js}] = \mathbb{E}[z'_{ijs} X_{is} \dot{X}^i_{js} \mathbb{E}[a_{js}|z_{ijs}, X_{is}, \dot{X}^i_{js}]] = 0$. Similarly, $\mathbb{E}[z'_{ijs} a_{js} \dot{X}^j_{is}] = \mathbb{E}[z'_{ijs} \dot{X}^j_{is} \mathbb{E}[a_{js}|z_{ijs}, \dot{X}^j_{is}] = 0$. Independence of a_{js} and a_{ks} when $k \neq j$ implies $\mathbb{E}[z'_{ijs} a_{js} a_{is} \dot{X}^i_{js}] = \mathbb{E}[z'_{ijs} \dot{X}^i_{js} \mathbb{E}[a_{js} a_{is}|z_{ijs}, \dot{X}^i_{js}] = 0$, and similarly $\mathbb{E}[z'_{ijs} a_{js} a_{is} \dot{a}^i_{js}] = 0$. Next,

$$\begin{aligned}\mathbb{E}[z_{ijs} a_{js} X_{is} \dot{a}^i_{js}] &= \mathbb{E}[z'_{ijs} a^2_{js} X_{is}] - \sum_{k\neq i} \mathbb{E}[z'_{ijs} a_{ks} X_{is}] \\ &= \frac{N-2}{N-1}\mathbb{E}[z'_{ijs} a^2_{js} X_{is}] - \sum_{k\neq i, k\neq j} \mathbb{E}[z'_{ijs} a_{ks} a_{js} X_{is}] \\ &= \frac{N-2}{N-1}\mathbb{E}[z'_{ijs} a^2_{js} X_{is}] \\ &= \frac{N-2}{N-1}\mathbb{E}[z'_{ijs} X_{is}]\sigma_a\end{aligned} \tag{A.53}$$

where σ_a is the variance of the scalar unobservable a. Similarly,

$$
\begin{aligned}
[z_{ijs}a_{js}\dot{a}^i_{js}] &= \mathbb{E}[z'_{ijs}a^2_{js}] - \sum_{k\neq i}\mathbb{E}[z'_{ijs}a_{ks}] \\
&= \frac{N-2}{N-1}\mathbb{E}[z'_{ijs}a^2_{js}] - \sum_{k\neq i,k\neq j}\mathbb{E}[z'_{ijs}a_{ks}a_{js}] \\
&= \frac{N-2}{N-1}\mathbb{E}[z'_{ijs}a^2_{js}] \\
&= \frac{N-2}{N-1}\mathbb{E}[z'_{ijs}]\sigma_a
\end{aligned}
\tag{A.54}
$$

Combining these results and substituting into Equation (A.52), now

$$
\begin{aligned}
\mathbb{E}[z'_{ijs}a_{is}\dot{g}^i_{ijs}] &= \mathbb{E}[z'_{ijs}a_{is}\dot{g}^i_{jis}]\tilde{\beta} + \frac{N-2}{N-1}\sigma_a(\mathbb{E}[z'_{ijs}X_{is}]\delta_2 + \mathbb{E}[z'_{ijs}]\gamma_2) \\
&= \mathbb{E}[z'_{ijs}a_{is}\dot{g}^i_{jis}]\tilde{\beta} + \frac{N-2}{N-1}\sigma_a\mathbb{E}[z'_{ijs}b^1_{ijs}]\begin{bmatrix}\delta_2 \\ \gamma_2\end{bmatrix}
\end{aligned}
\tag{A.55}
$$

Next, assume there exists $\theta_1 = (\tilde{\beta}, \delta_2, \gamma_2)$ and $\theta'_1 = (\tilde{\beta}', \delta'_2, \gamma'_2)$. Further, let σ^2_a and $(\sigma'_a)^2$ both be finite.

From Equation (A.55), it must be true that

$$
0 = \mathbb{E}[z'_{ijs}a_{is}\dot{g}^i_{jis}](\tilde{\beta} - \tilde{\beta}') + \frac{N-2}{N-1}\left(\sigma^2_a\mathbb{E}[z'_{ijs}b^1_{ijs}]\begin{bmatrix}\delta_2 \\ \gamma_2\end{bmatrix} - (\sigma'_a)^2\mathbb{E}[z'_{ijs}b^1_{ijs}]\begin{bmatrix}\delta'_2 \\ \gamma'_2\end{bmatrix}\right)
\tag{A.56}
$$

From above, $\tilde{\beta}$ is identified, and thus $(\tilde{\beta} - \tilde{\beta}') = 0$. Therefore,

$$
0 = \mathbb{E}[z'_{ijs}b^2_{ijs}]\left(\sigma^2_a\begin{bmatrix}\delta_2 \\ \gamma_2\end{bmatrix} - (\sigma_a)^2\begin{bmatrix}\delta'_2 \\ \gamma'_2\end{bmatrix}\right)
\tag{A.57}
$$

The second rank condition implies that there exists some $(m+1)\times l$ matrix $\mathbf{A_1}$ such that $\mathbf{A_1}\mathbb{E}[z'_{ijs}b^2_{ijs}]$ is of rank $2m$. Therefore, $(\mathbf{A_1}\mathbb{E}[z'_{ijs}b^1_{ijs}])^{-1}$ exists and

$$0 = \left(\sigma_a^2 \begin{bmatrix} \delta_2 \\ \gamma_2 \end{bmatrix} - (\sigma_a)^2 \begin{bmatrix} \delta'_2 \\ \gamma'_2 \end{bmatrix} \right) \tag{A.58}$$

Accordingly, δ_2 and γ_2 are identified up to the scale factor σ_a^2.

Step 2: Scale identification of δ_3 Multiply Equation (1.27) by $z'_{ijs}a_{is}$. So,

$$\begin{aligned} z'_{ijs}a_{is}\dot{g}^i_{ijs} &= z'_{ijs}a_{is}\dot{g}^i_{jis}\tilde{\beta} + z'_{ijs}a_{is}X_{is}\dot{X}^i_{js}\delta_1 + z'_{ijs}a_{is}X_{is}\dot{a}^i_{js}\delta_2 + z'_{ijs}a^2_{is}\dot{X}^i_{js}\delta_3 + z'_{ijs}a^2_{is}\dot{a}^i_{js}\delta_4 \\ &\quad + z'_{ijs}a_{is}\dot{X}^j_{is}\gamma_3 + z'_{ijs}a_{is}\dot{a}^i_{js}\gamma_4 - z'_{ijs}a_{is}\dot{c}^i_{ijs} \end{aligned} \tag{A.59}$$

Next, take the mean over all $SN(N-1)$ observations. So,

$$\begin{aligned} \frac{1}{SN(N-1)}\sum_{s=1}^{S}\sum_{i=1}^{N}\sum_{j\neq i} z'_{ijs}a_{is}\dot{g}^i_{ijs} &= \frac{1}{SN(N-1)}\sum_{s=1}^{S}\sum_{i=1}^{N}\sum_{j\neq i}(z'_{ijs}a_{is}\dot{g}^i_{jis}\tilde{\beta} + z'_{ijs}a_{is}X_{is}\dot{X}^i_{js}\delta_1 \\ &\quad + z'_{ijs}a_{is}X_{is}\dot{a}^i_{js}\delta_2 + z'_{ijs}a^2_{is}\dot{X}^i_{js}\delta_3 + z'_{ijs}a^2_{is}\dot{a}^i_{js}\delta_4 \\ &\quad + z'_{ijs}a_{is}\dot{X}^j_{is}\gamma_3 + z'_{ijs}a_{is}\dot{a}^i_{js}\gamma_4 - z'_{ijs}a_{is}\dot{c}^i_{ijs}) \end{aligned} \tag{A.60}$$

By the same argument as in Step 1, $\frac{1}{SN(N-1)}\sum_{s=1}^{S}\sum_{i=1}^{N}\sum_{j\neq i} z'_{ijs}a_{is}\dot{g}^i_{ijs} \to_p \mathbb{E}[z'_{ijs}a_{is}\dot{g}^i_{ijs}]$, etc. So, replace the matrices in Equation (A.60) with their probability limits.

$$\begin{aligned} \mathbb{E}[z'_{ijs}a_{is}\dot{g}^i_{ijs}] &= \mathbb{E}[z'_{ijs}a_{is}\dot{g}^i_{jis}]\tilde{\beta} + \mathbb{E}[z'_{ijs}a_{is}X_{is}\dot{X}^i_{js}]\delta_1 + \mathbb{E}[z'_{ijs}a_{is}X_{is}\dot{a}^i_{js}]\delta_2 + \mathbb{E}[z'_{ijs}a^2_{is}\dot{X}^i_{js}]\delta_3 \\ &\quad + \mathbb{E}[z'_{ijs}a^2_{is}\dot{a}^i_{js}]\delta_4 + \mathbb{E}[z'_{ijs}a_{is}\dot{X}^j_{is}]\gamma_3 + \mathbb{E}[z'_{ijs}a_{is}\dot{a}^i_{js}]\gamma_4 - \mathbb{E}[z'_{ijs}a_{is}\dot{c}^i_{ijs}] \end{aligned} \tag{A.61}$$

Assumption 1.6 implies $\mathbb{E}[z'_{ijs}a_{is}\dot{c}^i_{ijs}] = 0$. Further, $\mathbb{E}[z_{ijs}a_{is}X_{is}\dot{X}^i_{js}] = \mathbb{E}[z'_{ijs}X_{is}\dot{X}^i_{js}\mathbb{E}[a_{is}|z_{ijs}, X_{is}, \dot{X}^i_{js}]] = 0$ and similarly $\mathbb{E}[z_{ijs}a_{is}\dot{X}^i_{js}] = 0$. Independence of a_{is} and a_{js} from each other and from X implies $\mathbb{E}[z_{ijs}a_{is}X_{is}\dot{a}^i_{js}] = \mathbb{E}[z'_{ijs}X_{is}\mathbb{E}[\dot{a}^i_{js}\mathbb{E}[a_{is}|\dot{a}^i_{js}, X_{is}, z_{ijs}]|X_{is}, z_{ijs}]] = 0$, and by similar logic $\mathbb{E}[z'_{ijs}a_{is}\dot{a}_{js}] =$

0. Further,

$$
\begin{aligned}
[z'_{ijs}a^2_{is}\dot{a}^i_{js}] &= \mathbb{E}[z'_{ijs}a^2_{is}a_{js}] - \sum_{k\neq i}\mathbb{E}[z'_{ijs}a^2_{is}a_{ks}] \\
&= \mathbb{E}[z'_{ijs}a^2_{is}\mathbb{E}[a_{js}|a_{is}]] - \sum_{k\neq i}\mathbb{E}[z'_{ijs}a^2_{is}\mathbb{E}[a_{ks}|a_{is}]] \\
&= 0 \qquad \text{(A.62)}
\end{aligned}
$$

From Assumption 1.8, it follows that $\mathbb{E}[z'_{ijs}a^2_{is}\dot{X}^i_{js}] = \mathbb{E}[z'_{ijs}\dot{X}^i_{js}]\sigma^2_a$, where σ^2_a is the variance of a. Now, substittuion of these resutls into Equation (A.61) yields

$$\mathbb{E}[z'_{ijs}a_{is}\dot{g}^i_{ijs}] = \mathbb{E}[z'_{ijs}a_{is}\dot{g}^i_{jis}]\tilde{\beta} + \mathbb{E}[z'_{ijs}\dot{X}^i_{js}]\sigma^2_a\delta_3 \qquad \text{(A.63)}$$

Now, assume there exists some parameter vector $\theta_2 = (\tilde{\beta}, \gamma_3)$ and $\theta'_2 = (\tilde{\beta}', \gamma_3)$. These vectors are associated with finite σ^2_a and $(\sigma_a)^2$. So,

$$0 = \mathbb{E}[z'_{ijs}a_{is}\dot{g}^i_{jis}](\tilde{\beta} - \tilde{\beta}') + \mathbb{E}[z'_{ijs}\dot{X}^i_{js}](\sigma^2_a\delta_3 - (\sigma^2_a)'\delta'_3) \qquad \text{(A.64)}$$

Identification of $\tilde{\beta}$ implies $\tilde{\beta} = \tilde{\beta}'$. So,

$$0 = \mathbb{E}[z'_{ijs}\dot{X}^i_{js}](\sigma^2_a\delta_3 - (\sigma^2_a)'\delta'_3) \qquad \text{(A.65)}$$

The third rank condition further implies that there exists some mxl matrix $\mathbf{A_2}$ such that $\mathbf{A_2}\mathbb{E}[z'_{ijs}\dot{X}^i_{js}]$ is of full rank m. Therefore, $(\mathbf{A_2}\mathbb{E}[z'_{ijs}\dot{X}^i_{js}])^{-1}$ exists. So,

$$0 = \sigma^2_a\delta_3 - (\sigma^2_a)'\delta'_3 \qquad \text{(A.66)}$$

Accordingly, the parameter vector δ_3 is identified up to the scale factor σ^2_a.

Step 3: Scale identification of δ_4

Finally, multiply Equation (1.27) by $z'_{ijs}a_{is}a_{js}$. So,

$$z'_{ijs}a_{is}a_{js}\dot{g}^i_{ijs} = z'_{ijs}a_{is}a_{js}\dot{g}^i_{jis}\tilde{\beta} + z'_{ijs}a_{is}a_{js}X_{is}\dot{X}^i_{js}\delta_1 + z'_{ijs}a_{is}a_{js}X_{is}\dot{a}^i_{js}\delta_2 + z'_{ijs}a_{is}a^2_{js}\dot{X}^i_{js}\delta_3 \tag{A.67}$$

$$+ z'_{ijs}a^2_{is}a_{js}\dot{a}^i_{js}\delta_4 + z'_{ijs}a_{is}a_{js}\dot{X}^j_{is}\gamma_3 + z'_{ijs}a_{is}a_{js}\dot{a}^i_{js}\gamma_4 - z'_{ijs}a_{is}a_{js}\dot{c}^i_{ijs}$$

Next, take the mean over all $SN(N-1)$ observations. So,

$$\begin{aligned}\frac{1}{SN(N-1)}\sum_{s=1}^{S}\sum_{i=1}^{N}\sum_{j\neq i} z'_{ijs}a_{is}a_{js}\dot{g}^i_{ijs} = \frac{1}{SN(N-1)}\sum_{s=1}^{S}\sum_{i=1}^{N}\sum_{j\neq i}(&z'_{ijs}a_{is}a_{js}\dot{g}^i_{jis}\tilde{\beta} \\ &+ z'_{ijs}a_{is}a_{js}X_{is}\dot{X}^i_{js}\delta_1 \\ &+ z'_{ijs}a_{is}a_{js}X_{is}\dot{a}^i_{js}\delta_2 + z'_{ijs}a_{is}a^2_{js}\dot{X}^i_{js}\delta_3 + z'_{ijs}a^2_{is}a_{js}\dot{a}^i_{js}\delta_4 \\ &+ z'_{ijs}a_{is}a_{js}\dot{X}^j_{is}\gamma_3 + z'_{ijs}a_{is}a_{js}\dot{a}^i_{js}\gamma_4 - z'_{ijs}a_{is}a_{js}\dot{c}^i_{ijs})\end{aligned} \tag{A.68}$$

By the same argument as in Step 1, $\frac{1}{SN(N-1)}\sum_{s=1}^{S}\sum_{i=1}^{N}\sum_{j\neq i} z'_{ijs}a_{is}\dot{g}^i_{ijs} \to_p \mathbb{E}[z'_{ijs}a_{is}\dot{g}^i_{ijs}]$, etc. So, replace the matrices in Equation (A.68) with their probability limits, yielding

$$\begin{aligned}\mathbb{E}[z'_{ijs}a_{is}a_{js}\dot{g}^i_{ijs}] = {} & \mathbb{E}[z'_{ijs}a_{is}a_{js}\dot{g}^i_{jis}]\tilde{\beta} + \mathbb{E}[z'_{ijs}a_{is}a_{js}X_{is}\dot{X}^i_{js}]\delta_1 + \mathbb{E}[z'_{ijs}a_{is}a_{js}X_{is}\dot{a}^i_{js}]\delta_2 \\ & + \mathbb{E}[z'_{ijs}a_{is}a^2_{js}\dot{X}^i_{js}]\delta_3 + \mathbb{E}[z'_{ijs}a^2_{is}a_{js}\dot{a}^i_{js}]\delta_4 + \mathbb{E}[z'_{ijs}a_{is}a_{js}\dot{X}^j_{is}]\gamma_3 \\ & + \mathbb{E}[z'_{ijs}a_{is}a_{js}\dot{a}^i_{js}]\gamma_4 - \mathbb{E}[z'_{ijs}a_{is}a_{js}\dot{c}^i_{ijs}]\end{aligned} \tag{A.69}$$

Assumption 1.6 implies $\mathbb{E}[z'_{ijs}a_{is}a_{js}\dot{c}^i_{ijs}] = \mathbb{E}[z'_{ijs}a_{is}a_{js}\mathbb{E}[\dot{c}^i_{ijs}|z_{ijs}, a_{is}, a_{js}]] = 0$. Application of Assumption 1.6 and L.I.E. together imply $\mathbb{E}[z'_{ijs}a_{is}a_{js}X_{is}\dot{X}^i_{js}]$, $\mathbb{E}[z'_{ijs}a_{is}a_{js}X_{is}\dot{a}^i_{js}]$,

$\mathbb{E}[z'_{ijs}a_{is}a^2_{js}\dot{X}^i_{js}]$, $\mathbb{E}[z'_{ijs}a_{is}a_{js}\dot{X}^j_{is}]$, and $\mathbb{E}[z'_{ijs}a_{is}a_{js}\dot{a}^i_{js}]$ are also zero. Further,

$$\begin{aligned}
\mathbb{E}[z'_{ijs}a^2_{is}a_{js}\dot{a}^i_{js}] &= \mathbb{E}[z'_{ijs}a^2_{is}a^2_{js}] - \sum_{k\neq i}\mathbb{E}[z'_{ijs}a^2_{is}a_{js}a_{ks}] \\
&= \frac{N-2}{N-1}\mathbb{E}[z'_{ijs}a^2_{is}a^2_{js}] - \sum_{k\neq i,k\neq j}\mathbb{E}[z'_{ijs}a^2_{is}a_{js}a_{ks}] \\
&= \frac{N-2}{N-1}\mathbb{E}[z'_{ijs}a^2_{is}a^2_{js}] \\
&= \frac{N-2}{N-1}(\sigma^2_a)^2\mathbb{E}[z'_{ijs}]
\end{aligned} \tag{A.70}$$

Substitution into Equation (A.68) yields

$$\mathbb{E}[z'_{ijs}a_{is}a_{js}\dot{g}^i_{ijs}] = \mathbb{E}[z'_{ijs}a_{is}a_{js}\dot{g}^i_{jis}]\tilde{\beta} + \frac{n-2}{n-1}(\sigma^2_a)^2\mathbb{E}[z'_{ijs}]\delta_4 \tag{A.71}$$

Assume there exist parameter vectors $\theta_3 = (\tilde{\beta}, \delta_4)$ and $\theta'_3 = (\tilde{\beta}', \delta'_4)$, with associated σ^2_a and $(\sigma'_a)^2$. Equation (A.71) thus implies that

$$0 = \mathbb{E}[z'_{ijs}a_{is}a_{js}\dot{g}^i_{jis}](\tilde{\beta} - \tilde{\beta}') + \frac{n-2}{n-1}(\sigma^2_a)^2\mathbb{E}[z'_{ijs}](\delta_4 - \delta'_4) \tag{A.72}$$

Identification of β implies $(\tilde{\beta} - \tilde{\beta}') = 0$. Further, the fourth rank condition implies that there exists some $1 \times l$ matrix $\mathbf{A_3}$ such that $\mathbf{A_3}\mathbb{E}[z'_{ijs}]$ is of rank 1. Therefore, $0 = (\sigma^2_a)^2\delta_4 - ((\sigma'_a)^2)^2\delta'_4$, and δ_4 is identified to up to the scale factor σ^2_a.

□

A.8 Proposition 1.9

The prior propositions have provided conditions under which $\tilde{\beta}$, δ, and γ are identified. So, I proceed under the assumption that these parameters are identified. I now proceed to show that, conditional on these parameters being identified, a_{js} is identified for all j as $s \to \infty$.

First, for any i, j, k,

$$(\tilde{g}_{ijs} - \tilde{g}_{iks}) - \tilde{\beta}(\tilde{g}_{jis} - \tilde{g}_{kis}) = \delta_1 X_{is}(X_{js} - X_{ks}) + \delta_2 X_{is}(A_{js} - A_{ks}) + \delta_3 A_{is}(X_{js} - X_{ks}) + \delta_4 A_{is}(A_{js} - A_{ks}) + \gamma_3(X_{js} - X_{ks}) + \gamma_4(A_{js} - A_{ks}) - (\tilde{c}_{ijs} - \tilde{c}_{jis}) \tag{A.73}$$

Since every element on the right-hand side of Equation (A.73) is bounded, $(\tilde{g}_{ijs} - \tilde{g}_{iks}) - \tilde{\beta}(\tilde{g}_{jis} - \tilde{g}_{kis})$ is also bounded. Therefore, it has finite variance. Note further that it does not depend on N. Note that $\frac{1}{N-1}\sum_{k\neq i}\Big((\tilde{g}_{ijs} - \tilde{g}_{iks}) - \tilde{\beta}(\tilde{g}_{jis} - \tilde{g}_{kis})\Big) = \dot{g}^i_{ijs} - \tilde{\beta}\dot{g}^i_{jis}$.

Summing over $i \neq j$ and with slight rearrangement of Equation (1.27), for any j, we thus see

$$\frac{1}{(N-1)}\sum_{i\neq j}\left(\dot{g}^i_{ijs} - \tilde{\beta}\dot{g}^i_{jis}\right) = \frac{1}{(N-1)}\sum_{i\neq j}(\delta_1 X_{is}\dot{X}^i_{js} + \delta_2 X_{is}\dot{a}^i_{js} + \delta_3 a_{is}\dot{X}^i_{js} + \delta_4 a_{is}\dot{a}^i_{js} + \gamma_3 \dot{X}^j_{is} + \gamma_4 \dot{a}^i_{js} - \dot{c}^i_{ijs}) \tag{A.74}$$

Finite variance and independence implies that $\frac{1}{(N-1)}\sum_{i\neq j}\left(\dot{g}^i_{ijs} - \tilde{\beta}\dot{g}^i_{jis}\right) = \underset{i\neq j}{\mathbb{E}}[\left(\dot{g}^i_{ijs} - \tilde{\beta}\dot{g}^i_{jis}\right)] + o_p(1)$ for any j. Similarly,

- $\frac{1}{(N-1)}\sum_{i\neq j} X_{is}\dot{X}^i_{js} = X_{js}\underset{i\neq j}{\mathbb{E}}[X_{is}] - \underset{i\neq j}{\mathbb{E}}[X^2_{is}] + o_p(1)$
- $\frac{1}{(N-1)}\sum_{i\neq j} X_{is}\dot{a}^i_{js} = a_{js}\underset{i\neq j}{\mathbb{E}}[X_{is}] + o_p(1)$
- $\frac{1}{(N-1)}\sum_{i\neq j} a_{is}\dot{X}^i_{js} = o_p(1)$
- $\frac{1}{(N-1)}\sum_{i\neq j} a_{is}\dot{a}^i_{js} = o_p(1)$
- $\frac{1}{(N-1)}\sum_{i\neq j} \dot{X}^i_{js} = X_{js} - \underset{i\neq j}{\mathbb{E}}[X_{is}] + o_p(1)$
- $\frac{1}{(N-1)}\sum_{i\neq j} \dot{a}^i_{js} = a_{js} + o_p(1)$
- $\frac{1}{(N-1)}\sum_{i\neq j} \dot{c}^i_{ijs} = o_p(1)$

Therefore, in the limit, Equation (A.74) becomes

$$\underset{i \neq j}{\mathbb{E}}[\Big(\dot{g}^i_{ijs} - \tilde{\beta}\dot{g}^i_{jis}\Big)] = \delta_1(X_{js}\underset{i \neq j}{\mathbb{E}}[X_{is}] - \underset{i \neq j}{\mathbb{E}}[X^2_{is}]) + \delta_2 a_{js}\underset{i \neq j}{\mathbb{E}}[X_{is}]$$
$$+ \gamma_3(X_{js} - \underset{i \neq j}{\mathbb{E}}[X_{is}]) + \gamma_4 a_{js} + o_p(1) \tag{A.75}$$

Rearrangement yields

$$a_{js}(\gamma_4 + \delta_2\underset{i \neq j}{\mathbb{E}}[X_{is}]) = \underset{i \neq j}{\mathbb{E}}[\Big(\dot{g}^i_{ijs} - \tilde{\beta}\dot{g}^i_{jis}\Big)] - \delta_1(X_{js}\underset{i \neq j}{\mathbb{E}}[X_{is}] - \underset{i \neq j}{\mathbb{E}}[X^2_{is}])$$
$$- \gamma_3(X_{js} - \underset{i \neq j}{\mathbb{E}}[X_{is}]) + o_p(1) \tag{A.76}$$

Now, suppose there exist $a'_{js} \neq a_{js}$. From Equation (A.76), we see that $(a'_{js} - a_{js})(\gamma_4 + \delta_2\underset{i \neq j}{\mathbb{E}}[X_{is}]) = o_p(1)$. Therefore, $(\gamma_4 + \delta_2\underset{i \neq j}{\mathbb{E}}[X_{is}]) \neq 0 \Rightarrow (a'_{js} - a_{js}) = o_p(1)$ and thus a_{js} is point identified.

APPENDIX B

Supplementary Tables and Figures (Chapter 1)

Table B.1: Baseline Balance Across Schools

	Elected Treatment	Random Treatment	Control	P-value of Balance Test
Panel A: Baseline Covariates				
Elected	0.321	0.291	0.240	0.603
	(0.056)	(0.037)	(0.061)	
Grade 7	0.331	0.311	0.315	0.817
	(0.016)	(0.041)	(0.029)	
Grade 8	0.346	0.272	0.303	0.237
	(0.030)	(0.034)	(0.016)	
SC	0.195	0.267	0.285	0.572
	(0.057)	(0.068)	(0.084)	
ST	0.118	0.175	0.073	0.353
	(0.033)	(0.082)	(0.028)	
OBC	0.459	0.459	0.423	0.939
	(0.063)	(0.069)	(0.092)	
Panel B: Baseline Outcomes				
Education Aspirations	-0.217	-0.146	-0.201	0.892
	(0.144)	(0.105)	(0.096)	
Gender Roles	0.135	0.042	0.186	0.759
	(0.171)	(0.183)	(0.086)	

Notes: Robust standard errors in parentheses, clustered by school. Sample is 1319 students in 30 schools.

Table B.**2:** Baseline Balance Within Random Treatment Schools

	Participant	Non-Participant	P-value of Balance Test
Panel A: Baseline Covariates			
Elected	0.362	0.260	0.135
	(0.055)	(0.037)	
Grade 7	0.244	0.340	0.060
	(0.055)	(0.040)	
Grade 8	0.339	0.242	0.204
	(0.071)	(0.032)	
SC	0.283	0.260	0.779
	(0.084)	(0.074)	
ST	0.205	0.161	0.335
	(0.102)	(0.074)	
OBC	0.433	0.470	0.620
	(0.091)	(0.068)	
Panel B: Baseline Outcomes			
Educational Aspirations	-0.083	-0.174	0.489
	(0.142)	(0.111)	
Gender Roles	0.006	0.058	0.706
	(0.169)	(0.202)	

Notes: Robust standard errors in parentheses, clustered by school. Sample is 412 students in 10 Random Treatment schools.

APPENDIX C

Weighting in the Construction of Peer Means (Chapter 1)

In the standard setting with binary directed link data, peer weighting is a near-trivial matter and thus construction of peer means is fairly straightforward. In a binary setting, there are four obvious link definitions between individuals i and j:

1. An "OUT" link exists if individual i indicates that j is a friend.
2. An "IN" link exists if individual j indicates that i is a friend.
3. An "OR" link exists if either an "OUT" link or an "IN" link exists.
4. An "AND " link exists if both an "OUT" link and an "IN" link exist.

Note that the first two are necessarily directed, while the third and fourth are symmetric. For purposes of the reduced-form analysis in this paper, and to be consistent with the continuous results, I employ the "OUT" definition for binary network links.

Peer weighting is much more complicated when link intensities are continuous, as posited in the structural model developed in this paper. The following general assumptions on all weights will be maintained throughout. While in principle these weights could be estimated, in order to preserve computational power, I assume that the function is known. Letting g_{ijs} be the intensity of i's link toward j, and g_{jis} be the intensity of j's link toward i, the following three definitions seem natural

1. "OUT" link weight is g_{ijs}
2. An "IN" link weight is g_{jis}
3. An "SUM" link weight is $g_{ijs} + g_{jis}$

Once these weight are constructed, they are normalized so that the sum of the weights for a given individual i is one. For purposes of this paper, I employ the "SUM" weight definition for continuous link intensities. Future work will investigate the sensitivity of results to a choice of different weighting functions.

APPENDIX D

Graphical Reconstruction Algorithm (Chapter 1)

D.1 Network Data

Data imputation requires a model. As briefly discussed in the main body of text, given that I have a model of network formation, I use this model to impute missing network data. Imputation proceeds via an iterative EM algorithm. The algorithm proceeds as follows:

1. For the continuous network measure g_{ij}, impute missing data arbitrarily.

2. Using the imputed data, estimate the parameters of the network formation model. Recover moments of distributions of unobserved a_{is}, M_{is} and c_{ijs}

3. Using the implied distributions of the unobserved variables a_{is}, M_{is} and c_{ijs}, impute missing data. This step requires iteration of the network-formation process until an equilibrium consistent with the First-Order Conditions is reached.

4. Iterate Steps 2 and 3 sufficiently to reach convergence to a stable distribution of parameters and networks.

5. Take draws from this stable distribution. Construct point and variance estimates that properly adjust for imputation error. This adjustment is discussed in Cameron and Trivedi (2005) Section 27.7.

D.2 Outcome Data

Equation (1.7) provides the model's equation whereby outcomes are determined conditional on networks, observed variables, and unobserved a_{is}. Data imputation here proceeds from the imputed full networks as follows:

1. Take m draws from the converged distribution of networks and estimated parameters a_{is}.

2. For each draw

   2.1 Arbitrarily impute missing outcome data.

   2.2 Using imputed outcomes as well as the draw of networks and a_{is}, estimate the parameters of Equation (1.7).

   2.3 Using implied distribution of residuals from Step 2.2, impute outcome data where missing

   2.4 Iterate Steps 2.2 and 2.3 sufficient to reach convergence to stationary distribution. Take one draw from this distribution.

3. Given the final parameter values in Step 2.4, construct point and variance estimates that properly adjust for imputation error as well as error in estimating a_{is}.

APPENDIX E

Supplementary Tables (Chapter 2)

Table E.1: Summary of Simulated $\hat{\alpha}_1$

	$\gamma_1 = 0$			$\gamma_1 = -1$			$\gamma_1 = -2$		
$Pr(link)$	0.20	0.35	0.50	0.20	0.35	0.50	0.20	0.35	0.50
1	1.186	1.115	1.094	1.561	1.620	1.531	1.434	1.657	1.714
	[0.021]	[0.020]	[0.019]	[0.031]	[0.033]	[0.028]	[0.036]	[0.047]	[0.050]
2	1.186	1.115	1.094	1.287	1.424	1.406	1.189	1.401	1.531
	[0.020]	[0.019]	[0.019]	[0.026]	[0.033]	[0.031]	[0.030]	[0.043]	[0.050]
3	1.186	1.116	1.094	1.143	1.280	1.307	1.079	1.253	1.417
	[0.018]	[0.018]	[0.019]	[0.025]	[0.032]	[0.032]	[0.029]	[0.039]	[0.049]
4	1.186	1.116	1.095	1.070	1.165	1.221	1.023	1.141	1.324
	[0.018]	[0.018]	[0.018]	[0.025]	[0.032]	[0.033]	[0.029]	[0.039]	[0.048]
5	1.186	1.116	1.095	1.038	1.073	1.142	0.997	1.052	1.235
	[0.017]	[0.018]	[0.018]	[0.025]	[0.033]	[0.034]	[0.030]	[0.038]	[0.048]
6	1.186	1.116	1.094	1.027	1.003	1.069	0.986	0.982	1.147
	[0.017]	[0.017]	[0.018]	[0.026]	[0.033]	[0.035]	[0.030]	[0.039]	[0.049]
7	1.186	1.116	1.094	1.024	0.956	1.001	0.983	0.930	1.054
	[0.017]	[0.017]	[0.017]	[0.026]	[0.035]	[0.037]	[0.030]	[0.040]	[0.050]
8	1.186	1.116	1.094	1.023	0.929	0.942	0.982	0.899	0.963
	[0.017]	[0.017]	[0.017]	[0.026]	[0.035]	[0.038]	[0.030]	[0.040]	[0.052]
9	1.186	1.116	1.094	1.023	0.916	0.893	0.982	0.884	0.880
	[0.017]	[0.017]	[0.017]	[0.026]	[0.035]	[0.039]	[0.030]	[0.041]	[0.054]
10	1.186	1.116	1.094	1.023	0.910	0.857	0.982	0.878	0.814
	[0.017]	[0.017]	[0.017]	[0.026]	[0.036]	[0.039]	[0.030]	[0.041]	[0.056]
11	1.186	1.116	1.094	1.023	0.908	0.834	0.982	0.876	0.771
	[0.017]	[0.017]	[0.017]	[0.026]	[0.036]	[0.040]	[0.030]	[0.041]	[0.058]
12	1.186	1.116	1.094	1.023	0.908	0.822	0.982	0.876	0.750
	[0.017]	[0.017]	[0.016]	[0.026]	[0.036]	[0.040]	[0.030]	[0.041]	[0.058]
13	1.186	1.116	1.094	1.023	0.908	0.817	0.982	0.875	0.743
	[0.017]	[0.017]	[0.016]	[0.026]	[0.036]	[0.040]	[0.030]	[0.041]	[0.059]
14	1.186	1.116	1.094	1.023	0.908	0.815	0.982	0.875	0.743
	[0.017]	[0.017]	[0.016]	[0.026]	[0.036]	[0.040]	[0.030]	[0.041]	[0.058]
15	1.186	1.116	1.094	1.023	0.908	0.814	0.982	0.876	0.744
	[0.017]	[0.017]	[0.016]	[0.026]	[0.036]	[0.040]	[0.030]	[0.041]	[0.058]
16	1.186	1.116	1.094	1.023	0.908	0.814	0.982	0.876	0.745
	[0.017]	[0.017]	[0.016]	[0.026]	[0.036]	[0.040]	[0.030]	[0.041]	[0.058]
17	1.186	1.116	1.094	1.023	0.908	0.814	0.982	0.876	0.745
	[0.017]	[0.017]	[0.016]	[0.026]	[0.036]	[0.040]	[0.030]	[0.041]	[0.058]
18	1.186	1.116	1.094	1.023	0.908	0.814	0.982	0.876	0.745
	[0.017]	[0.017]	[0.016]	[0.026]	[0.036]	[0.040]	[0.030]	[0.041]	[0.058]
19	1.186	1.116	1.094	1.023	0.908	0.814	0.982	0.876	0.745
	[0.017]	[0.017]	[0.016]	[0.026]	[0.036]	[0.040]	[0.030]	[0.041]	[0.058]
20+	1.186	1.116	1.094	1.023	0.908	0.814	0.982	0.876	0.745
	[0.017]	[0.017]	[0.016]	[0.026]	[0.036]	[0.040]	[0.030]	[0.041]	[0.058]

Notes: Table presents mean simulated estimates. Standard deviations of simulated estimates in brackets. 1000 simulations performed with 200 schools of 25 students each (N = 5000).

Table E.2: Summary of Simulated $\hat{\alpha}_2$

	$\gamma_1 = 0$			$\gamma_1 = -1$			$\gamma_1 = -2$		
$Pr(link)$	0.20	0.35	0.50	0.20	0.35	0.50	0.20	0.35	0.50
1	0.477	0.265	0.205	0.954	0.470	0.266	1.604	0.974	0.527
	[0.026]	[0.025]	[0.024]	[0.043]	[0.036]	[0.033]	[0.048]	[0.054]	[0.052]
2	0.899	0.522	0.405	1.476	0.811	0.476	2.004	1.341	0.774
	[0.037]	[0.037]	[0.037]	[0.052]	[0.048]	[0.044]	[0.051]	[0.059]	[0.060]
3	1.197	0.765	0.596	1.796	1.099	0.670	2.214	1.584	0.952
	[0.041]	[0.048]	[0.048]	[0.055]	[0.056]	[0.052]	[0.053]	[0.062]	[0.066]
4	1.360	0.995	0.782	1.978	1.358	0.859	2.333	1.784	1.112
	[0.043]	[0.058]	[0.059]	[0.057]	[0.062]	[0.060]	[0.054]	[0.065]	[0.069]
5	1.427	1.203	0.963	2.064	1.585	1.053	2.394	1.956	1.274
	[0.045]	[0.064]	[0.068]	[0.058]	[0.068]	[0.067]	[0.054]	[0.067]	[0.074]
6	1.448	1.376	1.140	2.097	1.771	1.251	2.419	2.100	1.445
	[0.045]	[0.068]	[0.076]	[0.059]	[0.071]	[0.075]	[0.055]	[0.069]	[0.079]
7	1.453	1.500	1.309	2.107	1.904	1.449	2.428	2.209	1.629
	[0.045]	[0.071]	[0.083]	[0.059]	[0.075]	[0.082]	[0.055]	[0.071]	[0.085]
8	1.454	1.575	1.466	2.110	1.987	1.639	2.431	2.279	1.816
	[0.045]	[0.072]	[0.088]	[0.059]	[0.077]	[0.088]	[0.055]	[0.071]	[0.089]
9	1.455	1.611	1.604	2.110	2.030	1.808	2.431	2.317	1.993
	[0.045]	[0.074]	[0.092]	[0.059]	[0.078]	[0.095]	[0.055]	[0.072]	[0.094]
10	1.455	1.625	1.711	2.110	2.050	1.942	2.431	2.334	2.141
	[0.045]	[0.074]	[0.094]	[0.059]	[0.079]	[0.098]	[0.055]	[0.073]	[0.098]
11	1.455	1.629	1.783	2.110	2.057	2.036	2.431	2.340	2.245
	[0.045]	[0.074]	[0.096]	[0.059]	[0.079]	[0.101]	[0.055]	[0.073]	[0.100]
12	1.455	1.630	1.823	2.110	2.059	2.094	2.431	2.342	2.305
	[0.045]	[0.074]	[0.097]	[0.059]	[0.079]	[0.103]	[0.055]	[0.073]	[0.102]
13	1.455	1.631	1.842	2.110	2.059	2.123	2.431	2.343	2.332
	[0.045]	[0.074]	[0.097]	[0.059]	[0.079]	[0.104]	[0.055]	[0.073]	[0.103]
14	1.455	1.631	1.848	2.110	2.059	2.136	2.431	2.343	2.341
	[0.045]	[0.074]	[0.097]	[0.059]	[0.079]	[0.104]	[0.055]	[0.073]	[0.102]
15	1.455	1.631	1.850	2.110	2.059	2.140	2.431	2.343	2.343
	[0.045]	[0.074]	[0.097]	[0.059]	[0.079]	[0.105]	[0.055]	[0.073]	[0.102]
16	1.455	1.631	1.851	2.110	2.059	2.142	2.431	2.343	2.343
	[0.045]	[0.074]	[0.097]	[0.059]	[0.079]	[0.105]	[0.055]	[0.073]	[0.102]
17	1.455	1.631	1.851	2.110	2.059	2.142	2.431	2.343	2.343
	[0.045]	[0.074]	[0.097]	[0.059]	[0.079]	[0.104]	[0.055]	[0.073]	[0.102]
18	1.455	1.631	1.851	2.110	2.059	2.142	2.431	2.343	2.343
	[0.045]	[0.074]	[0.097]	[0.059]	[0.079]	[0.105]	[0.055]	[0.073]	[0.102]
19	1.455	1.631	1.851	2.110	2.059	2.142	2.431	2.343	2.343
	[0.045]	[0.074]	[0.097]	[0.059]	[0.079]	[0.105]	[0.055]	[0.073]	[0.102]
20+	1.455	1.631	1.851	2.110	2.059	2.142	2.431	2.343	2.343
	[0.045]	[0.074]	[0.097]	[0.059]	[0.079]	[0.105]	[0.055]	[0.073]	[0.102]

Notes: Table presents mean simulated estimates. Standard deviations of simulated estimates in brackets. 1000 simulations performed with 200 schools of 25 students each (N = 5000).

Table E.3: Summary of Simulated $\hat{\beta}_1$

	$\gamma_1 = 0$			$\gamma_1 = -1$			$\gamma_1 = -2$		
$Pr(link)$	0.20	0.35	0.50	0.20	0.35	0.50	0.20	0.35	0.50
1	0.449	0.598	0.728	0.870	0.787	0.717	1.384	1.822	1.938
	[0.071]	[0.112]	[0.132]	[0.097]	[0.173]	[0.274]	[0.157]	[0.335]	[0.675]
2	0.591	0.726	0.822	0.684	0.696	0.747	0.824	0.986	1.004
	[0.045]	[0.060]	[0.066]	[0.053]	[0.088]	[0.110]	[0.070]	[0.127]	[0.196]
3	0.619	0.747	0.839	0.654	0.702	0.772	0.700	0.785	0.807
	[0.035]	[0.044]	[0.046]	[0.043]	[0.061]	[0.077]	[0.050]	[0.078]	[0.115]
4	0.613	0.734	0.834	0.633	0.709	0.786	0.655	0.725	0.741
	[0.031]	[0.037]	[0.038]	[0.037]	[0.049]	[0.056]	[0.044]	[0.058]	[0.081]
5	0.605	0.704	0.817	0.615	0.700	0.788	0.628	0.705	0.728
	[0.029]	[0.034]	[0.033]	[0.036]	[0.042]	[0.046]	[0.042]	[0.047]	[0.064]
6	0.601	0.669	0.790	0.604	0.681	0.782	0.612	0.689	0.733
	[0.029]	[0.033]	[0.031]	[0.036]	[0.038]	[0.040]	[0.041]	[0.043]	[0.051]
7	0.600	0.639	0.758	0.600	0.655	0.767	0.604	0.668	0.741
	[0.029]	[0.033]	[0.030]	[0.036]	[0.037]	[0.036]	[0.041]	[0.041]	[0.044]
8	0.599	0.617	0.721	0.598	0.631	0.742	0.602	0.643	0.741
	[0.029]	[0.034]	[0.031]	[0.036]	[0.037]	[0.034]	[0.041]	[0.041]	[0.039]
9	0.599	0.605	0.683	0.598	0.614	0.711	0.601	0.622	0.728
	[0.029]	[0.034]	[0.032]	[0.036]	[0.038]	[0.034]	[0.041]	[0.040]	[0.038]
10	0.599	0.600	0.650	0.598	0.605	0.678	0.601	0.609	0.701
	[0.029]	[0.034]	[0.034]	[0.036]	[0.038]	[0.035]	[0.041]	[0.041]	[0.038]
11	0.599	0.598	0.625	0.598	0.601	0.648	0.601	0.602	0.668
	[0.029]	[0.034]	[0.035]	[0.036]	[0.039]	[0.037]	[0.041]	[0.041]	[0.040]
12	0.599	0.597	0.609	0.598	0.599	0.625	0.601	0.599	0.638
	[0.029]	[0.034]	[0.036]	[0.036]	[0.039]	[0.038]	[0.041]	[0.041]	[0.041]
13	0.599	0.597	0.602	0.598	0.599	0.610	0.601	0.599	0.617
	[0.029]	[0.034]	[0.036]	[0.036]	[0.039]	[0.039]	[0.041]	[0.041]	[0.042]
14	0.599	0.597	0.599	0.598	0.599	0.603	0.601	0.598	0.605
	[0.029]	[0.034]	[0.036]	[0.036]	[0.039]	[0.039]	[0.041]	[0.041]	[0.043]
15	0.599	0.597	0.598	0.598	0.599	0.599	0.601	0.598	0.600
	[0.029]	[0.034]	[0.036]	[0.036]	[0.039]	[0.039]	[0.041]	[0.041]	[0.044]
16	0.599	0.597	0.598	0.598	0.599	0.598	0.601	0.598	0.598
	[0.029]	[0.034]	[0.036]	[0.036]	[0.039]	[0.039]	[0.041]	[0.041]	[0.044]
17	0.599	0.597	0.598	0.598	0.599	0.598	0.601	0.598	0.597
	[0.029]	[0.034]	[0.036]	[0.036]	[0.039]	[0.039]	[0.041]	[0.041]	[0.044]
18	0.599	0.597	0.598	0.598	0.599	0.598	0.601	0.598	0.597
	[0.029]	[0.034]	[0.036]	[0.036]	[0.039]	[0.039]	[0.041]	[0.041]	[0.044]
19	0.599	0.597	0.598	0.598	0.599	0.598	0.601	0.598	0.597
	[0.029]	[0.034]	[0.036]	[0.036]	[0.039]	[0.039]	[0.041]	[0.041]	[0.044]
20+	0.599	0.597	0.598	0.598	0.599	0.598	0.601	0.598	0.597
	[0.029]	[0.034]	[0.036]	[0.036]	[0.039]	[0.039]	[0.041]	[0.041]	[0.044]

Notes: Table presents mean simulated estimates. "True" $\beta_1 = 0.6$. Standard deviations of simulated estimates in brackets. 1000 simulations performed with 200 schools of 25 students each (N = 5000).

Table E.4: Summary of Simulated $\hat{\beta}_2$

	$\gamma_1 = 0$			$\gamma_1 = -1$			$\gamma_1 = -2$		
$Pr(link)$	0.20	0.35	0.50	0.20	0.35	0.50	0.20	0.35	0.50
1	1.038	0.996	0.979	1.338	1.496	1.468	1.219	1.443	1.589
	[0.029]	[0.026]	[0.024]	[0.035]	[0.042]	[0.036]	[0.033]	[0.053]	[0.067]
2	0.992	0.970	0.964	1.176	1.360	1.381	1.132	1.361	1.531
	[0.021]	[0.019]	[0.018]	[0.026]	[0.034]	[0.034]	[0.030]	[0.045]	[0.056]
3	0.987	0.966	0.961	1.076	1.254	1.318	1.062	1.264	1.457
	[0.019]	[0.017]	[0.017]	[0.024]	[0.032]	[0.035]	[0.028]	[0.040]	[0.052]
4	0.992	0.969	0.961	1.028	1.174	1.267	1.025	1.186	1.395
	[0.017]	[0.016]	[0.016]	[0.023]	[0.032]	[0.033]	[0.028]	[0.038]	[0.050]
5	0.997	0.975	0.964	1.009	1.112	1.222	1.008	1.124	1.339
	[0.017]	[0.016]	[0.015]	[0.023]	[0.031]	[0.033]	[0.028]	[0.037]	[0.048]
6	0.999	0.983	0.968	1.003	1.066	1.180	1.002	1.076	1.285
	[0.017]	[0.016]	[0.015]	[0.023]	[0.030]	[0.033]	[0.028]	[0.037]	[0.048]
7	1.000	0.990	0.973	1.002	1.034	1.141	1.000	1.041	1.230
	[0.017]	[0.016]	[0.015]	[0.023]	[0.030]	[0.032]	[0.028]	[0.037]	[0.047]
8	1.000	0.995	0.980	1.001	1.016	1.104	1.000	1.019	1.175
	[0.017]	[0.016]	[0.015]	[0.023]	[0.030]	[0.031]	[0.028]	[0.036]	[0.047]
9	1.000	0.997	0.986	1.001	1.006	1.071	1.000	1.007	1.122
	[0.017]	[0.016]	[0.015]	[0.023]	[0.030]	[0.031]	[0.028]	[0.036]	[0.046]
10	1.000	0.999	0.991	1.001	1.002	1.044	1.000	1.002	1.075
	[0.017]	[0.016]	[0.015]	[0.023]	[0.030]	[0.030]	[0.028]	[0.036]	[0.046]
11	1.000	0.999	0.996	1.001	1.000	1.024	1.000	1.000	1.038
	[0.017]	[0.016]	[0.015]	[0.023]	[0.030]	[0.030]	[0.028]	[0.036]	[0.046]
12	1.000	0.999	0.998	1.001	1.000	1.011	1.000	0.999	1.015
	[0.017]	[0.016]	[0.015]	[0.023]	[0.030]	[0.030]	[0.028]	[0.036]	[0.046]
13	1.000	0.999	1.000	1.001	1.000	1.005	1.000	0.999	1.004
	[0.017]	[0.016]	[0.015]	[0.023]	[0.030]	[0.030]	[0.028]	[0.036]	[0.046]
14	1.000	0.999	1.000	1.001	1.000	1.002	1.000	0.999	1.000
	[0.017]	[0.016]	[0.015]	[0.023]	[0.030]	[0.030]	[0.028]	[0.036]	[0.045]
15	1.000	0.999	1.000	1.001	1.000	1.001	1.000	0.999	0.999
	[0.017]	[0.016]	[0.015]	[0.023]	[0.030]	[0.030]	[0.028]	[0.036]	[0.045]
16	1.000	0.999	1.000	1.001	1.000	1.001	1.000	0.999	0.999
	[0.017]	[0.016]	[0.015]	[0.023]	[0.030]	[0.030]	[0.028]	[0.036]	[0.045]
17	1.000	0.999	1.000	1.001	1.000	1.001	1.000	0.999	0.999
	[0.017]	[0.016]	[0.015]	[0.023]	[0.030]	[0.030]	[0.028]	[0.036]	[0.045]
18	1.000	0.999	1.000	1.001	1.000	1.001	1.000	0.999	0.999
	[0.017]	[0.016]	[0.015]	[0.023]	[0.030]	[0.030]	[0.028]	[0.036]	[0.045]
19	1.000	0.999	1.000	1.001	1.000	1.001	1.000	0.999	0.999
	[0.017]	[0.016]	[0.015]	[0.023]	[0.030]	[0.030]	[0.028]	[0.036]	[0.045]
20+	1.000	0.999	1.000	1.001	1.000	1.001	1.000	0.999	0.999
	[0.017]	[0.016]	[0.015]	[0.023]	[0.030]	[0.030]	[0.028]	[0.036]	[0.045]

Table presents mean simulated estimates. Notes: "True" $\beta_2 = 1$. Standard deviations of simulated estimates in brackets. 1000 simulations performed with 200 schools of 25 students each (N = 5000).

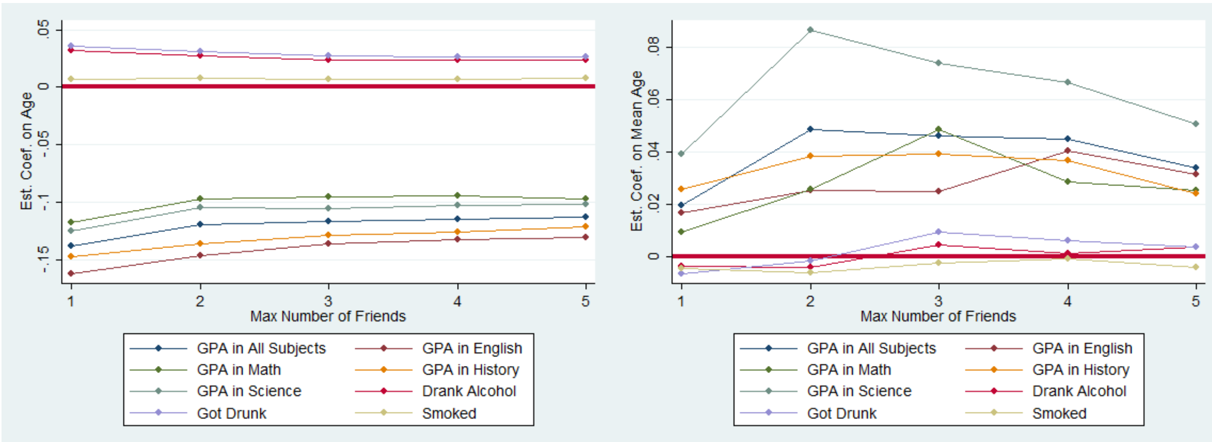
Table E.5: Summary of Simulated $\hat{\beta}_3$

	$\gamma_1 = 0$			$\gamma_1 = -1$			$\gamma_1 = -2$		
$Pr(link)$	0.20	0.35	0.50	0.20	0.35	0.50	0.20	0.35	0.50
1	-0.077	-0.421	-0.614	-1.004	-1.053	-0.960	-2.405	-3.631	-3.706
	[0.093]	[0.136]	[0.158]	[0.236]	[0.350]	[0.483]	[0.482]	[0.885]	[1.520]
2	0.082	-0.410	-0.631	-0.181	-0.613	-0.872	-0.493	-1.260	-1.505
	[0.070]	[0.091]	[0.096]	[0.137]	[0.193]	[0.214]	[0.224]	[0.355]	[0.462]
3	0.275	-0.290	-0.569	0.132	-0.412	-0.794	0.034	-0.554	-0.941
	[0.061]	[0.079]	[0.082]	[0.115]	[0.146]	[0.167]	[0.162]	[0.227]	[0.283]
4	0.409	-0.132	-0.481	0.320	-0.241	-0.706	0.259	-0.247	-0.681
	[0.057]	[0.074]	[0.078]	[0.102]	[0.128]	[0.132]	[0.145]	[0.175]	[0.210]
5	0.473	0.044	-0.374	0.428	-0.062	-0.596	0.387	-0.069	-0.543
	[0.056]	[0.073]	[0.076]	[0.100]	[0.117]	[0.120]	[0.140]	[0.149]	[0.177]
6	0.494	0.211	-0.247	0.478	0.109	-0.468	0.455	0.077	-0.448
	[0.055]	[0.075]	[0.078]	[0.100]	[0.111]	[0.113]	[0.136]	[0.139]	[0.152]
7	0.500	0.344	-0.107	0.497	0.262	-0.322	0.485	0.213	-0.352
	[0.056]	[0.074]	[0.078]	[0.100]	[0.108]	[0.109]	[0.137]	[0.136]	[0.141]
8	0.501	0.431	0.041	0.502	0.376	-0.155	0.495	0.333	-0.235
	[0.056]	[0.075]	[0.083]	[0.100]	[0.109]	[0.107]	[0.137]	[0.134]	[0.135]
9	0.501	0.476	0.188	0.503	0.446	0.017	0.498	0.420	-0.089
	[0.056]	[0.075]	[0.087]	[0.100]	[0.110]	[0.109]	[0.136]	[0.133]	[0.134]
10	0.501	0.494	0.317	0.503	0.482	0.178	0.499	0.470	0.079
	[0.056]	[0.075]	[0.091]	[0.100]	[0.112]	[0.111]	[0.136]	[0.134]	[0.135]
11	0.501	0.500	0.409	0.503	0.497	0.311	0.499	0.493	0.240
	[0.056]	[0.075]	[0.093]	[0.100]	[0.112]	[0.116]	[0.136]	[0.135]	[0.140]
12	0.501	0.502	0.465	0.503	0.501	0.404	0.499	0.502	0.364
	[0.056]	[0.076]	[0.095]	[0.100]	[0.112]	[0.118]	[0.136]	[0.135]	[0.143]
13	0.501	0.502	0.491	0.503	0.503	0.459	0.499	0.505	0.442
	[0.056]	[0.076]	[0.095]	[0.100]	[0.112]	[0.118]	[0.136]	[0.135]	[0.145]
14	0.501	0.502	0.501	0.503	0.503	0.487	0.499	0.505	0.483
	[0.056]	[0.076]	[0.096]	[0.100]	[0.112]	[0.119]	[0.136]	[0.135]	[0.147]
15	0.501	0.502	0.504	0.503	0.503	0.498	0.499	0.505	0.499
	[0.056]	[0.076]	[0.096]	[0.100]	[0.112]	[0.119]	[0.136]	[0.135]	[0.148]
16	0.501	0.502	0.505	0.503	0.503	0.502	0.499	0.505	0.505
	[0.056]	[0.076]	[0.096]	[0.100]	[0.112]	[0.119]	[0.136]	[0.135]	[0.148]
17	0.501	0.502	0.505	0.503	0.503	0.503	0.499	0.505	0.507
	[0.056]	[0.076]	[0.096]	[0.100]	[0.112]	[0.119]	[0.136]	[0.135]	[0.148]
18	0.501	0.502	0.505	0.503	0.503	0.503	0.499	0.505	0.507
	[0.056]	[0.076]	[0.096]	[0.100]	[0.112]	[0.119]	[0.136]	[0.135]	[0.148]
19	0.501	0.502	0.505	0.503	0.503	0.503	0.499	0.505	0.507
	[0.056]	[0.076]	[0.096]	[0.100]	[0.112]	[0.119]	[0.136]	[0.135]	[0.148]
20+	0.501	0.502	0.505	0.503	0.503	0.503	0.499	0.505	0.507
	[0.056]	[0.076]	[0.096]	[0.100]	[0.112]	[0.119]	[0.136]	[0.135]	[0.148]

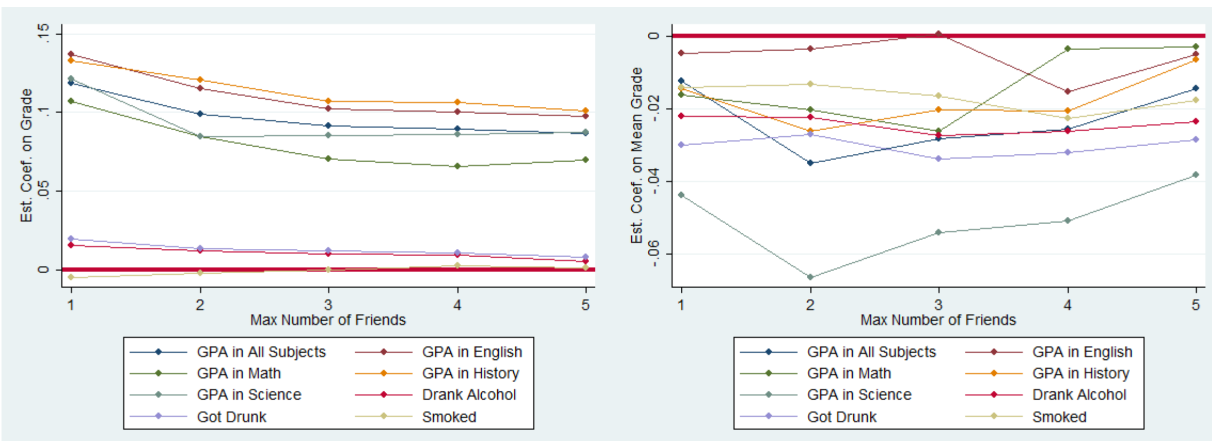
Notes: Table presents mean simulated estimates. "True" $\beta_3 = 0.5$. Standard deviations of simulated estimates in brackets. 1000 simulations performed with 200 schools of 25 students each (N = 5000).

Figure E.1: Add Health Estimated Coefficients (Estimator 2)

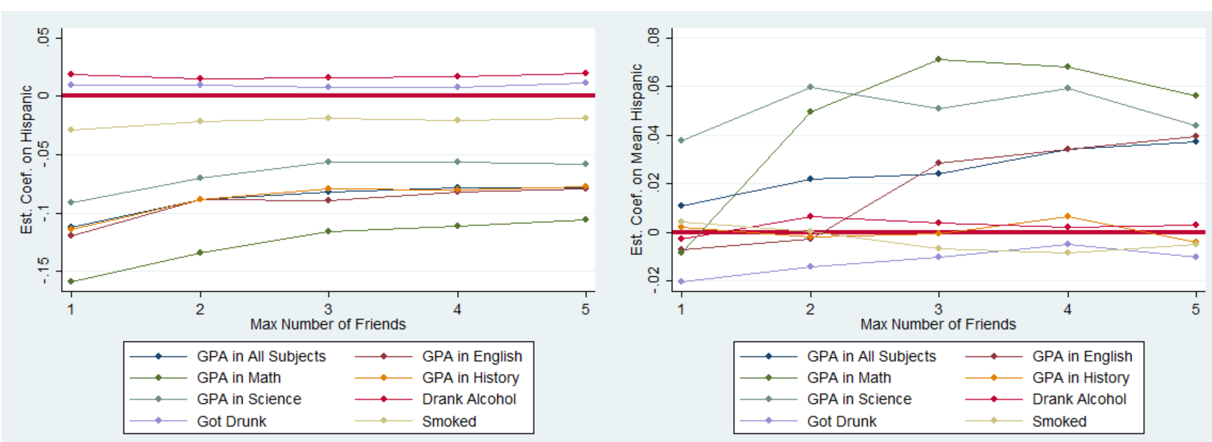
(a) Coefficients for Age and $\overline{\text{Age}}$



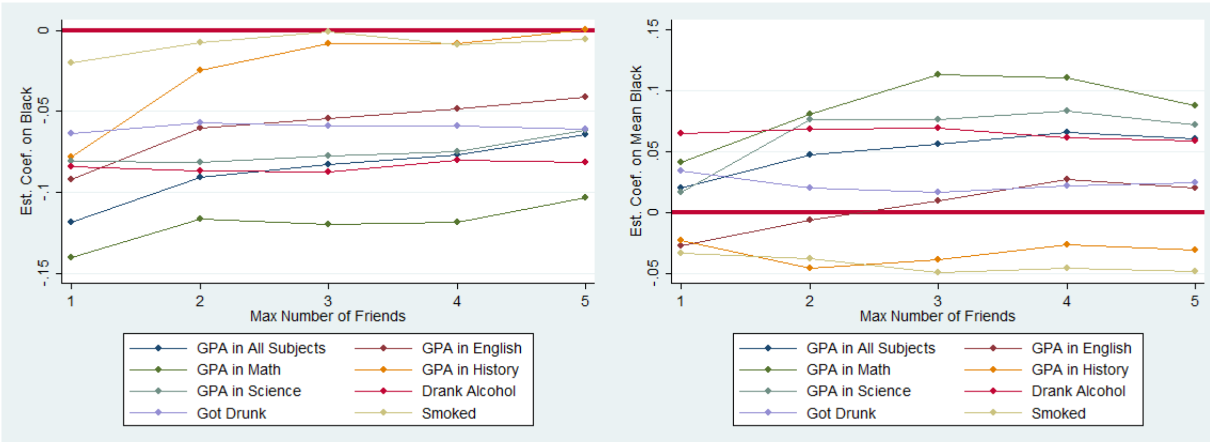
(b) Coefficients for Grade and $\overline{\text{Grade}}$



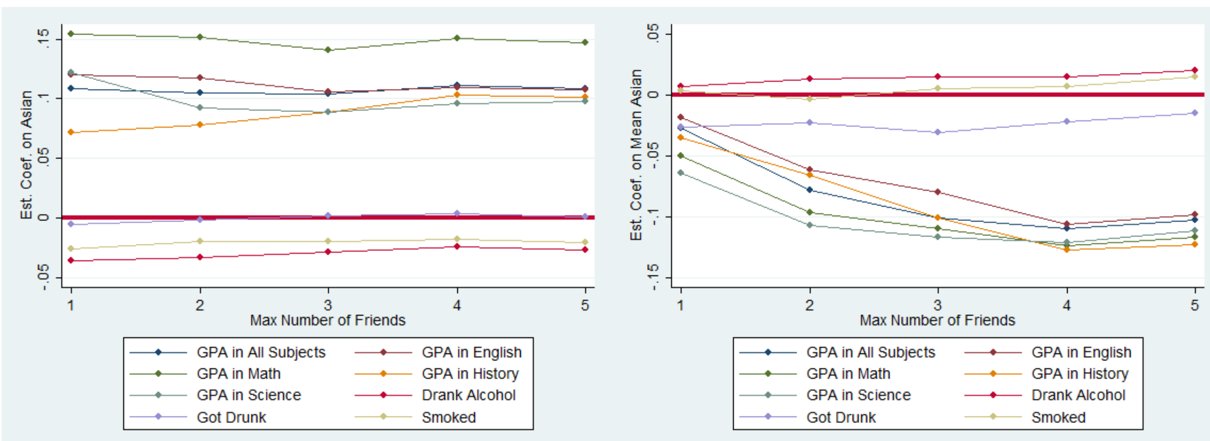
(c) Coefficients for Hispanic and $\overline{\text{Hispanic}}$



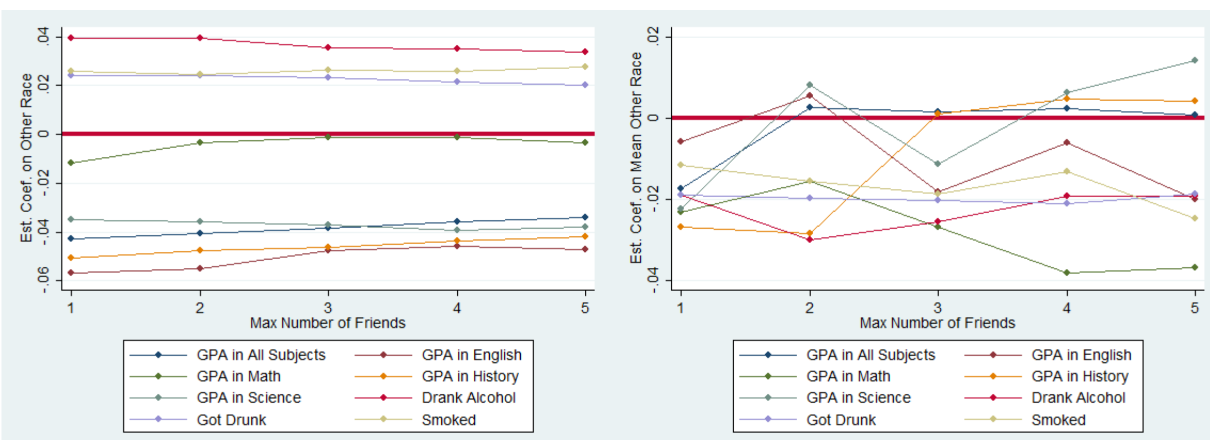
(d) Coefficients for Black and $\overline{\text{Black}}$



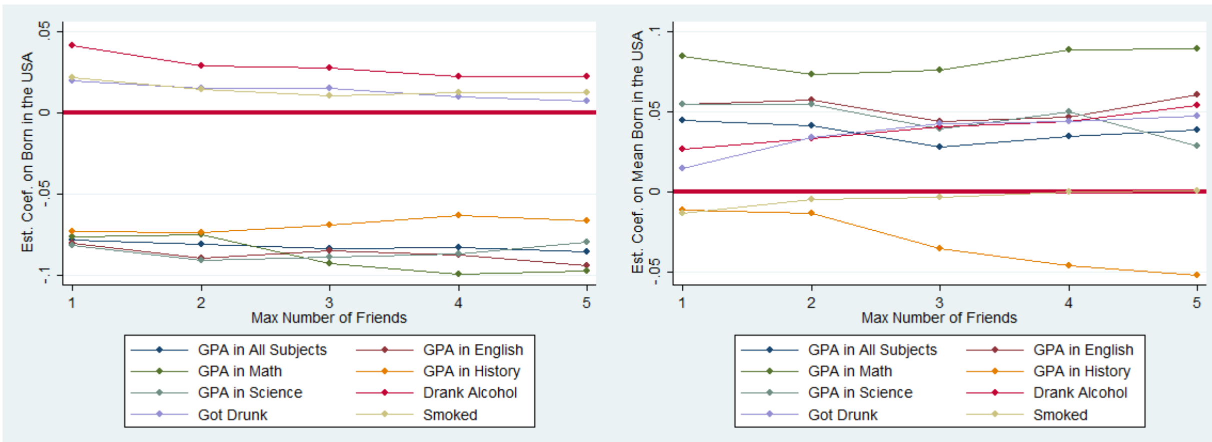
(e) Coefficients for Asian and $\overline{\text{Asian}}$



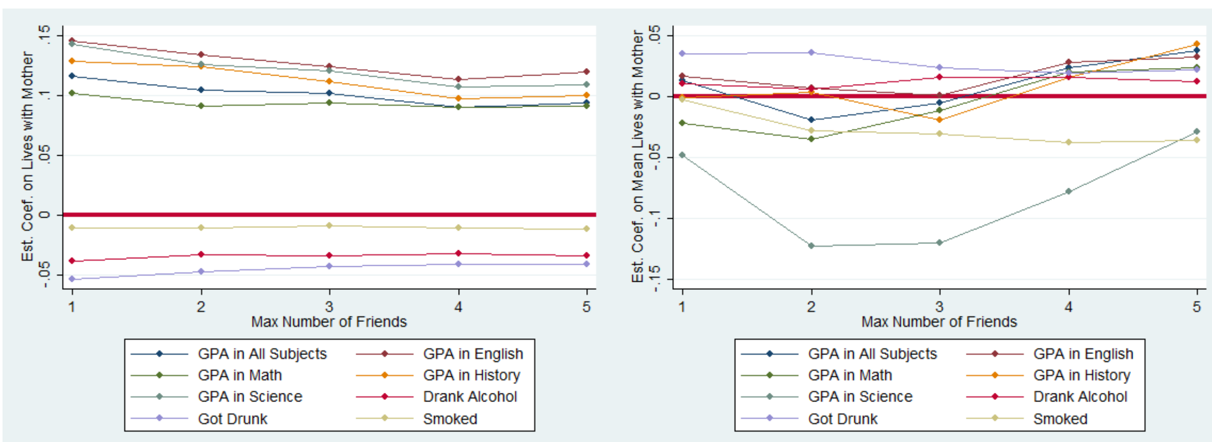
(f) Coefficients for Other Race and $\overline{\text{Other Race}}$



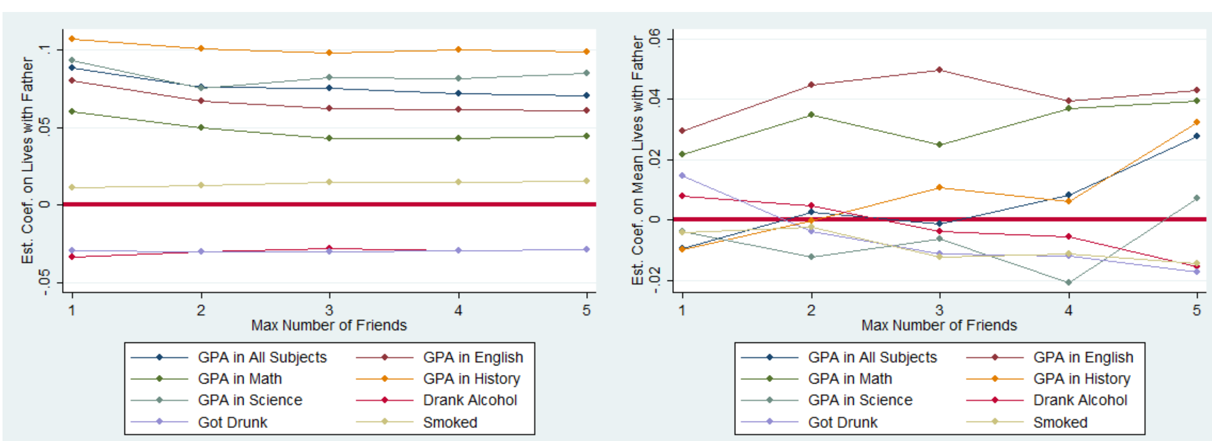
(g) Coefficients for Born in the USA and $\overline{\text{Born in the USA}}$



(h) Coefficients for Lives with Mother and $\overline{\text{Lives wtih Mother}}$



(i) Coefficients for Lives with Father and $\overline{\text{Lives with Father}}$



APPENDIX F

Supplementary Tables (Chapter 3)

Table F.1: Differential Attrition

Panel A: Differential Attrition by Treatment Arm

Present for	Baseline Network (1)	Endline Qnr (2)	Endline Network (3)	Baseline Network (4)	Endline Qnr (5)	Endline Network (6)
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 Backward 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: Differential Attrition Across Participation Status within T2

Present for	Baseline Network (1)	Endline Qnr (2)	Endline Network (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 Backward 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
Gemder Roles Index	0.801	0.695	0.901

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 Baseline Qnr (N = 1,053 girls, 894 boys).
Sample for Panel B is all girls present for Baseline Qnr in T2 schools (N = 338).
Panel A presents P-values of a test that the coefficient on the two treatment dummies interacted with baseline characteristics are jointly significant.
Panel A presents P-values of a test that the coefficient on "Selected in T2" interacted with baseline characteristics is significant.

BIBLIOGRAPHY

Abadie, Alberto, Matthew M. Chingos, and Martin R. West. 2014. "Endogenous Stratification in Randomized Experiments." Unpublished Working Paper.

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer. 2015. "Identification Properties of Recent Production Function Estimators." *Econometrica*, 83(6): 2411–2451.

Altermatt, Ellen Rydell, and Evan M. Pomerantz. 2003. "The Development of Competence-Related and Motivational Beliefs." *Journal of Educational Psychology*, 95(1): 111.

Angrist, Joshua. 2014. "The Perils of Peer Effects." *Labour Economics*, 30: 98–108.

Attanasio, Orazio P., Costas Meghir, and Ana Santiago. 2011. "Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate PROGRESA." *Review of Economic Studies*, 79: 37–66.

Badev, Anton. 2013. "Discrete Games in Endogenous Networks: Theory and Policy." Unpublished Working Paper.

Bailey, Richard. 2005. "Evaluating the Relationship between Physical Education, Sport and Social Inclusion." *Educational Review*, 57(1): 71–90.

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

Banerjee, Abhijit, Arun G. Chandrasekhar, Ester Duflo, and Matthew O. Jackson. 2012. "The Diffusion of Microfinance." NBER Working Paper No. 17743.

Baumann, Leonie. 2016. "Time Allocation in Friendship Networks." Unpublished Working Paper.

Beaman, Lori, Ester Duflo, and Rohini Pande adn Petia Topalova. 2012. "Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India." *Science*, 335(6068): 582–586.

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

Bearman, Peter S., and James Moody. 2004. "Suicide and Friendships among American Sdolescents." *American Journal of Public Health*, 94(1): 89–95.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Authomobile Prices and Market Equilibrium." *Econometrica*, 63(4): 841–890.

Bhattacharya, Debopam. 2009. "Inferring Optimal Peer Assignment from Experimental Data." *Journal of the American Statistical Association*, 104(486): 486–500.

Bhattacharya, Debopam, and Pascaline Dupas. 2012. "Inferring Welfare Maximizing Treatment Assignment under Budget Constraints." *Journal of Econometrics*, 167: 168–196.

Black, Dan A., and Jeffrey A. Smith. 2006. "Estimating the Returns to College Quality with Multiple Proxies for Quality." *Journal of Labor Economics*, 24(3): 701–728.

Bloch, Francis, and Bhaskar Dutta. 2009. "Communication Networks with Endogenous Link Strength." *Games and Economic Behavior*, 66(1): 39–56.

Blume, Lawrence E., William A. Brock, Steven N. Durlauf, and Rajshri Jayaraman. 2015. "Linear Social Interactions Models." *Journal of Political Economy*, 123(2): 444–496.

Booij, Adam, Edwin Leuven, and Hessel Oosterbeek. 2016. "Ability Effects in University: Evidence from a Randomized Experiment." *Review of Economic Studies*. Forthcoming.

Boucher, Vincent. 2015. "Structural Homophily." *International Economic Review*, 56(1): 235–264.

Boucher, Vincent, Yann Bramoullé, Habiba Djebbari, and Bernard Fortin. 2014. "Do Peers Affect Student Achievement? Evidence from Canada Using Group Size Variation." *Journal of Applied Econometrics*, 29(1): 91–109.

Bound, John, Charles Brown, and Nancy Mathiowetz. 2001. "Measurement Error in Survey Data." In *Handbook of Econometrics*. Vol. 5, 3805–3843.

Bramoullé, Yann. 2013. "Comment." *Journal of Business and Economic Statistics*, 31(3): 264–266.

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin. 2009. "Identification of Peer Effects through Social Networks." *Journal of Econometrics*, 150(1): 41–55.

Burgess, Simon, and Marcela Umana-Aponte. 2011. *Raising Your Sights: The Impact of Friendship Aspirations on Educational Aspirations.* Centre for Market Organization.

Burgess, Simon, Eleanor Sanderson, and Marcela Omana Aponte. 2011. "School Ties: An Analysis of Homophily in an Adolescent Friendship Network." The Centre for Market and Public Organisation. Department of Economics, University of Briston, UK.

Cai, Jing, Alain De Janvry, and Elisabeth Sadoulet. 2015. "Social Networks and the Decision to Insure." *American Economic Journal: Applied Economics*, 7(2): 81–108.

Cameron, A. Colin, and Pravin K. Trivedi. 2005. *Microeconometrics: Methods and Applications.* New York:Cambridge University Press.

Carrell, Scott, Bruce Sacerdote, and James West. 2013. "From Random Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation." *Econometrica*, 81(3): 855–882.

Carrell, Scott, Mark Hoekstra, and James West. 2011. "Is Poor Fitness Contagious? Evidence from Randomly Assigned Friends." *Journal of Public Economics*, 95: 657–663.

Carrell, Scott, Richard Fullerton, and James West. 2009. "Does Your Cohort Matter? Measuring Peer Effects in College Achievement." *Journal of Labor Economics*, 27(3): 439–464.

Chandrasekhar, Arun. 2015. "Econometrics of Network Formation." In *Oxford Handbook on the Econometrics of Networks.*

Chandrasekhar, Arun, and Matthew O. Jackson. 2014. "Tractable and Consistent Random Graph Models." NBER Working Paper No. 20276.

Chandrasekhar, Arun, and Randall Lewis. 2011. "Econometrics of Sampled Networks." Unpublished Working Paper.

Christakis, Nicholas A., and James H. Fowler. 2007. "The Spread of Obesity in a Large Social Network over 32 Years." *New England Journal of Medicine*, 357(4): 370–379.

Christakis, Nicholas, James H. Fowler, Guido W. Imbens, and Karthik Kalyanaraman. 2010. "An Empirical Model of Strategic Network Formation." NBER Working Paper 16039.

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

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

Cunha, Flavio, James J. Heckman, and Susanne M. Shennach. 2010. "Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Econometrica*, 78(3): 883–931.

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

DeGiorgi, Giacomo, Michele Pelllizzari, and Silvia Radaelli. 2010. "Identification of Social Interactions through Partially Overlapping Peer Groups." *American Economic Journal: Applied Economics*, 2: 241–255.

Dehejia, Rajeev H. 2005. “Program Evaluation as a Decision Problem.” *Journal of Econometrics*, 125: 141–173.

Delavallade, Clara, Alan Griffith, and Rebecca Thornton. 2016. “Unintended Spillovers of a Girls’ Empowerment Program in Rural Rajasthan: Targeting Matters.” Unpublished Working Paper.

de Paula, Áureo, Seth Richards-Shubik, and Elie Tamer. 2016. “Identifcation of Preferences in Network Formation Games.” CENMAP Working Paper CWP29/15.

Dercon, Stefan, and Abhijeet Singh. 2013. “From Nurition to Aspirations and Self-Sufficiency: Gender Bias over Time among Children in Four Countries.” *World Development*, 45: 31–50.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2011. “Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya.” *American Economic Review*, 101(5): 1739–1774.

Duflo, Esther, Rema Hanna, and Stephen P. Ryan. 2012. “Incentives Work: Getting Teachers to Come to School.” *American Economic Review*, 102(4): 1241–1278.

Epple, Dennis, and Richard Romano. 2011. “Peer Effects in Education: A Survey of the Theory and Evidence.” In *Handbook of Social Economics*. Vol. 1, 1053–1165.

Fiorina, Morris P. 1999. “Extreme Voices: A Dark Side of Civic Engagement.” In *Civic Engagement in American Democracy*. Vol. 395, 405–413.

French, D., E. Jansen, M. Riansari, and K. Setiono. 2003. “Friendships of Indonesian Children: Adjustment of Children Who Differ in Friendship Presence and SImilarity between Mutual Friends.” *Social Development*, 12(4): 605–621.

Goldsmith-Pinkham, Paul, and Guido Imbens. 2013. “Social Networks and the Identification of Peer Effects.” *Journal of Business and Economic Statistics*, 31(3): 253–264.

Graham, Bryan S. 2014*a*. “An Econometric Model of Link Formation with Degree Heterogeneity.” NBER Working Paper No. 20341.

Graham, Bryan S. 2014*b*. “Methods of Identification in Social Networks.” NBER Working Paper No. 20414.

Griffith, Alan. 2017. “Random Assignmetn with Non-Random Peers: A Structural Approach to Counterfactual Treatment Assessment.” Unpublished Working Paper.

Hamm, Jill V. 2000. “Do Birds of a Feather Flock Together? The Variable Bases for African American, Asican American, and European American Adolescents’ Selection of Similar Friends.” *Developmental Psychology*, 36(2): 209–219.

Harris, Kathleen Mullan. 2009. *The National Longitudinal Study of Adolescent to Adult Health (Add Health), Waves I and II, 1994-1996.* Carolina Population Center, University of North Carolina at Chapel Hill.

Heckman, James, and Edward Vytlacil. 1998. "Instrumental Variables Methods for the Correlated Random Effects Coefficient Model." *Journal of Human Resources*, 33(4): 974–987.

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

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica*, 76(3): 643–660.

Hsieh, Chih-Sheng, and Lung Fei Lee. 2016. "A Social Interactions Model with Endogenous Friendship Formation and Selectivity." *Journal of Applied Econometrics*, 31: 301–319.

Jackson, Matthew O. 2008. *Social and Economic Networks.* Princeton, New Jersey:Princeton Press.

Jackson, Matthew O. 2014. "Networks in the Understanding of Human Behavior." *Journal of Economic Perspectives*, 28(4): 3–22.

Kandel, Denise B. 1978. "Homophily, Selection, and Socialization in Adolescent Friendships." *American Journal of Sociology*, 84(2): 427–436.

Kandpal, Eeshani, and Kathy Baylis. 2013. "Expanding Horizons: Can Women's Support Groups Diversify Peer Networks in Rural India?" *American Journal of Agricultural Economics*, 95(2): 360–367.

Kazianga, Harounan, Dan Levy, and Leign L. Linden. 2013. "The Effects of "Girl-Friendly" Schools: Evidence from the BRIGHT Construction Program in Burkina Faso." *American Economic Journal: Applied Economics*, 41–62.

Lee, Lung-fei. 2007. "Identification and Estimation of Econometric Models with Group Interactions, Contextual Factors, and Fixed Effects." *Journal of Econometrics*, 140(2): 333–374.

Leung, Michael. 2015. "Two-Step Estimation of Network-Formation Models with Incomplete Information." *Journal of Econometrics*, 188(1): 182–195.

Levinsohn, James, and Amil Petrin. 2003. "Estimating Production funcitons Using Inputs to Control for Unobservables." *Review of Economic Studies*, 70(2): 317–341.

Lyle, David S. 2009. "The Effects of Peer Group Heterogeneity on the Production of Human Capital at West Point." *American Economic Jouranl: Applied Economics*, 1(4): 69–84.

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

Manski, Charles F. 2004. "Statistical Treatment Rules for Heterogeneous Populations." *Econometrica*, 72(4): 1221–1246.

Mele, Angelo. 2010. "A Structural Model of Segregation in Social Networks." Unpublished Working Paper.

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

Monderer, Dov, and Lloyd S. Shapley. 1996. "Potential Games." *Games and Economic Behavior*, 14(1): 124–143.

Nations, United. 2015. "Transforming Our World: The 2030 Agenda for Sustainable Development." Resolution adopted by the General Assembly on 25 September 2015, A/RES/70/1.

Ngatia, Muthoni. 2015. "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.

Olley, G. Steven, and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Industry." *Econometrica*, 64(4): 1263–1297.

Oster, Emily, and Rebecca Thornton. 2011. "Menstruation, Sanitary Products, and School Attendance: Evidence from a Randomized Evaluation." *American Economic Journal: Applied Economics*, 3(1): 91–100.

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.

Rogers, Brian W. 2006. "Learning and Status in Social Networks." Ph.D. Dissertation. California Institute of Technology.

Rosen, J. Ben. 1965. "Existence and Uniqueness of Equilibrium Points for Concave N-Person Games." *Econometrica*, 520–534.

Rubin, Donald B. 1976. "Inference and Missing Data." *Biometrika*, 581–592.

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

Sacerdote, Bruce. 2011. "Peer Effects in Education: How Might They Work, How Big Are They, and How Much Do We Know Thus Far?" In *Hanbook of the Economics of Education*. Vol. 3, 249–277.

Shadish, William R., M.H. Clark, and Peter M. Steiner. 2008. "Can Nonrandomized Experiments Yield Accurate Answers? A Randomized Experiment Comparing Random and Nonrandom Assignments." *Journal of the American Statistical Association*, 103: 1334–1354.

Sheng, Shuyang. 2012. "Identification and Estimation of Network Formation Games." Unpublished Working Paper.

Shortall, Sally. 2008. "Are Rural Development Programmes Socially Inclusive? Social Inclusion, Civic Engagement, Participation and Social Capital: Exploring the Differences." *Journal of Rural Studies*, 24(4): 450–457.

Smith, Jeffrey, and Jonas Staghøj. 2009. "Using Statistical Treatment Rules for Assignment of Participants in Labor Market Programs." Unpublished Working Paper.

Sojourner, Aaron. 2013. "Identification of Peer Effects with Missing Peer Data: Evidence from Project STAR." *Economic Journal*, 123(569): 574–605.

Stinebrickner, Ralph, and Todd Stinebrickner. 2006. "What Can Be Learned about Peer Effects Using College Roommates? Evidence from New Survey Data and Students from Disadvantaged Backgrounds." *Journal of Public Economics*, 90: 1435–1454.

Todd, Petra, and Kenneth Wolpin. 2006. "Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility." *American Economic Review*, 96(5): 1384–1414.

Ui, Takashi. 2008. "Correlated Equilibrium and Concave Games." *International Journal of Game Theory*, 37(1): 1–13.

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.

Williams, Kevin M. 2016. "Social Networks in Higher Education: Estimating Spillovers with Experimental Variation." Unpublished Working Paper.

Wooldridge, Jeffrey M. 1997. "On Two Stage Least Squares Estimation of the Average Treatment Effect in a Random Coefficient Model." *Economics Letters*, 56: 129–133.

Wooldridge, Jeffrey M. 2003. "Further Results on Instrumental Variables Estimation of Average Treatment Effects in the Correlated Random Coefficient Model." *Economics Letters*, 79: 185–191.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data.* Cambridge, Massachusetts:MIT Press.