Essays on Peer Effects in Social Networks

By

Tam Bui

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Economics

May 31st, 2020

Nashville, Tennessee

Approved:

Tong Li, Ph.D., Chair

Andrew Dustan, Ph.D.

Atsushi Inoue, Ph.D.

Pedro Sant'Anna, Ph.D.

Lijun Song, Ph.D

To my parents

Acknowledgments

Above all, I want to express my sincere gratitude to my advisor, Professor Tong Li, for all of his guidance and support with such kindness and generosity throughout my time at Vanderbilt. It is my great opportune to be his student. I want to thank my committee members for their valuable advice and feedback on this work. I wish to acknowledge the love of my family, my parents and sister, who are my pillar of strength. I am indebted to my friends in and outside of the department, in alphabetical order, Alper Arslan, Phuong Bui, Thu Tran Coleman, Huong Dao, Linh Dao, Bich Hoang, Trang Hoang, Trang Luu, Daniel Mangrum, Linh Nguyen, Chi Pham, Nam Pham, and Jun Zhao; whose warmth, humour, intellect, and maturity, have kept me afloat.

This dissertation uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth). No direct support was received from grant P01-HD31921 for this analysis.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction: Social Networks and Peer Effects

The study of peer effects has captured the attention of researchers in many fields.[1] Policymakers are interested in how a development program influences not only its participants but also the participants' friends, relatives, and co-workers. Media corporations, marketers, and politicians are invested in learning how to select an optimal group of influencers that spread news, ideas, and opinions quickly and at a low cost.

Obtaining reliable estimates of peer effects, however, is a challenge. First and foremost, social networks exhibit homophily, i.e. people with similar characteristics are likely to be friends with each other (Lazarsfeld and Merton, 1954; McPherson et al., 2001). Do we observe the true influence of peers' behavior on one's choice, or only that similar people are likely to choose similar things in the first place? Researchers have taken various strategies to deal with this complication. One is to find believable scenarios where network connections can be seen as sufficiently random (e.g. Sacerdote, 2001; Angrist and Lang, 2004; Carrell et al., 2009; Imberman et al., 2012; Carrell et al., 2013), while another is to incorporate the network formation process (e.g. Goldsmith-Pinkham and Imbens, 2013; Hsieh and Lee, 2016; Badev, 2017; Liu et al., 2017).[2]

Another obstacle to the estimation of peer effects is the accuracy of network information. There might be network members and network connections that are not observed in the data, which can distort network characteristics in general (Lee et al., 2006; Chandrasekhar and Lewis, 2016; Hsieh et al., 2019). To complicate the matter further, social connections are likely to be dynamic in the sense that a network changes over time or manifests differently in various contexts. The connections providing advice on taking a loan could be different from the friends who recommend a new novel. It is still rare to

[1]To follow the terms coined by Manski (1993), we can think of peer effects in three types: 'endogenous' peer effects (friends'outcomes affect one's outcome), contextual peer effects (friends' exogenous characteristics affect one's outcome), and correlated effects (friends face correlated shocks due to common environment and context, or could be a result of endogenous network formation based on unobserved individual characteristics of members). We are most interested in the identification and estimation of 'endogenous' peer effects

[2]Of course, network formation in itself helps answer other important questions: How do social structures emerge and sustain over time? Can we remove barriers between groups that human beings create to separate ourselves?

obtain data that can capture this fluidity of network.[3] Without a more accurate description of the network structure that is relevant to the question of interest, estimation of peer effects is strenuous.

This dissertation brings into focus situations where researchers may observe settings about which they have an idea how the structure of peer effects may change dynamically, and changes in an endogenous way. Consider the classic analysis of peer effects of students' academic achievements, now with a twist: GPAs within a specific course. If students have some freedom in course selection, the resulting connections can be categorized into two groups: those who take the same course and those who do not. It is difficult to argue that subject GPAs of philosophy students are equally influenced if at all by the unrealized subject GPAs of friends who have never taken philosophy courses, ones whose potentials remain unknown and unobservable to both their peers and researchers. Researchers want to assume an arguably more realistic network structure where only students attending the same course can influence each other. Yet, one difficulty arises as they recognize students may be more likely to choose a course in which they may perform well. As a result, this hypothesized structure of networks should not be assumed to be random. In this situation, a traditional model of network formation is not directly applicable because, again, we do not actually observe the exact network manifestation.

In order to analyze this setting, in Chapter 2 we formalize a model that captures the dynamics and selection in social networks described above, then we propose a two-step estimation method for the model to obtain better estimates of peer effects. We show the consistency of the procedure and suggest block bootstrapping procedure for inference. In Chapter 3 we apply the estimation to a real data set to estimate peer effects among high school students' subject GPAs. While the correction for sample selection seems not to significantly influence the estimates of peer effects, assumptions about how network structures for different types of peer effects change have strong implications on the estimates.

[3] There are surveys that collect detailed information about the type of peers related to specific outcomes. Such attention is valuable, yet outcomes of interest do not always align as planned and researchers need to adjust network appropriately to the problem at hand.

Chapter 2: Sample Selection with Peer Effects: Modelling and Estimation Method

2.1 Introduction

In this chapter, we focus on the modeling and estimation of peer effects taking into account the fluidity of networks due to sample selection as motivated in the introduction. We abstract away from network formation and instead, we think of the original network (who is friends with whom) as a fixed and exogenous starting point but the structure of peer effects (who influences whom for which outcomes) as dynamic and endogenous that evolves from the original network and the choices made earlier.[1] To use the same motivating example as before, for instance, we assume that students' choices to take courses in these subjects are correlated with their subject GPAs. To add to this, the original network can affect students' selections via the influence of friends' selection probabilities. Once students have taken the classes in the subject, they are then also influenced directly by friends' subject GPAs. However, the type of peers that matters for GPAs now differs, as a result of choices at the beginning of the semester. Specifically, we assume that students' subject performance is influenced only by that of friends who are also taking courses in the same subject at the same time, and not by the past or future performance of peers who ever did or will take courses in the subject.

We propose a two-stage sample selection model with peer effects present in both stages to characterize this process of realizing subject-specific GPAs. The subject selection decision in the first period is modeled by a binary choice under incomplete information. The subject-specific GPA realization is characterized by the linear spatial autoregressive (SAR) model. The possible correlation between the random components in two stages poses a form of endogenous network structure, while the missing of some outcomes can be seen as a type of sample selection. In this chapter, we prove consistency of a two-stage estimator that corrects for bias via a control function approach. We evaluate the performance of bootstrap to make inference for this two-stage estimation procedure via

[1]The network structure in our model is dynamic in the sense that it can be shaped inadvertently by choices of members in the network. It is not dynamic in the sense that members actively create and remove connections, which requires the data on the new network.

Monte Carlo simulation.

Several specifications of network models with sample selection have been considered in the spatial and network econometrics literature. Wang and Lee (2013) considers the case where the outcome variable is missing at random. Hoshino (2017) considers missing outcome variable by selection but assumes that there are no peer effects in the selection process. The incorporation of dependence in the selection process while providing a friendly estimator proves to be a challenge. McMillen (1995); Flores-Lagunes and Schnier (2012); Rabovic and Cizek (2016) consider processes where there is dependence in both selection and outcome stages. These papers model the selection process by different non-linear spatial variants of the spatial autoregressive model and the spatial autoregressive error model. Estimation of these spatial probit models involves high dimensional integration. McMillen (1995) develops an expectation-maximization (EM) algorithm. However, the estimator is infeasible and computationally intensive and a rigorous theory is not developed. Flores-Lagunes and Schnier (2012) uses results by Pinkse and Slade (1998) shows that it is possible estimate with GMM, by taking into account the heteroskedasticity from spatial correlation and neglecting spatial autocorrelation. At the same time, these models are concerned with the case where network structures in any stage, selection and outcome, are exogenous. The key difference in our model is that the selection decision in the first period in fact shapes the network later on. The use of nonlinear spatial models to characterize the first stage selection process as in the spatial econometrics literature will be non-tractable. We show that employing a strategic binary choice under incomplete information is attractive and still captures the dependence characteristic of the selection stage.

2.2 Model

Before specifying the model, we first define some notations. Let d_i denote the student's selection decision in the first period, y_i^* denote the student's subject-specific GPA realized in the second period. $\mathbf{z}_i$ and $\mathbf{x}_i$ are the row vectors of students' exogenous characteristics. We require that $\mathbf{x}_i$ is a subset of $\mathbf{z}_i$. v_i and u_i are the idiosyncratic errors. The column vector and matrix notations are $\mathbf{d}_n$, $\mathbf{y}^*_n$, $\mathbf{Z}_n$, $\mathbf{X}_n$, $\mathbf{v}_n$, and $\mathbf{u}_n$.

We characterize the friendship among students by an $n \times n$ adjacency matrix G_n where its element $g_{ij} \in (0, 1)$ denotes whether student i nominates j as a friend. By convention, $g_{ii} = 0$. In our model, we do not require friendship to be reciprocal so that it may be the case $g_{ij} \neq g_{ji}$. For example, consider the relationships among four individuals as in the following Fig. 2.1:

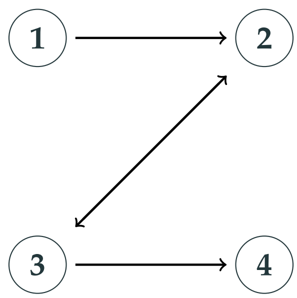


Figure 2.1: Example of a simple network structure

An adjacency matrix describing relations in Fig. 2.1 can be written as:

$$G = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Let $\mathbf{g}_i = (g_{i1}, g_{i2}, \ldots, g_{in})$ be the $1 \times n$ row vector specifying all the nominations of i to everyone else in the network of n individuals. We call G_n the original raw network structure. G^*_n is the row-normalized version of G_n such that if student i nominates anyone as a friend the ith row sum is equal to 1 and 0 otherwise. The corresponding elements of

G_n^* are g_{ij}^* and $\mathbf{g}_i^*$.[2] The normalized adjacency matrix for Fig. 2.1 can be written as:

$$G^* = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

In our setting, researchers can observe the full network structure G_n, the covariate matrices $\mathbf{Z}_n$, $\mathbf{X}_n$ as well as the subject selection of all students in the first period $\mathbf{d}_n$. However, we only observe a subset of subject-specific GPAs for students who actually take any course in the subject so that $y_n = D_n \mathbf{y}_n^*$ for $D_n = \text{diag}(\mathbf{d}_n)$.

2.2.1 Binary selection stage and incomplete information

We assume that students' subject selection decision d_i given the selection of other people in the network $\mathbf{d}_{-i}$ can be characterized as to maximize a reduced-form latent utility function:

$$U_i^s(d_i = 1, \mathbf{d}_{-i}) = \mathbf{z}_i\gamma + \delta \sum_j g_{ij}^* d_j + v_i = \mathbf{z}_i\gamma + \delta \mathbf{g}_i^* \mathbf{d}_n \tag{2.2.1}$$

$$U_i^s(d_i = 0, \mathbf{d}_{-i}) = 0$$

for v_i be iid across individuals. The term $\delta \mathbf{g}_i^* \mathbf{d}_n$ can be seen as the extra utility student i gains from having some proportion of friends who also take the same subject. δ captures the peer effects that the average friends' selection has on student i and is therefore a parameter of main interest.[3]

[2] In the general SAR framework, an interaction matrix is not necessarily row-normalized. Some theoretical models may suggest a local aggregate model in which the 'raw' interaction matrix G_n is used, so that individuals is influenced by the sum of friends' outcomes instead of average of friends' outcome (see Liu et al. (2014)), which still assumes a type of homogenous influence.

[3] Manski (1993) coins several terms for different types of peer effects: endogenous peer effect (friends' outcomes affect one's outcome), contextual peer effect (friends' exogenous characteristics affect one's outcome), and correlated effect (friends face correlated shocks due to common environment and context). We are mostly concerned with 'endogenous peer effects' and use the term 'peer effects' interchangeably with 'endogenous peer effects'. When contextual peer effects are added, we will use the full name to distinguish.

Suppose students make decisions about taking courses in a subject each semester simultaneously and likely to be in a short period of time, we assume that students do not have perfect knowledge about friends' actual selection.[4] We define an incomplete information setting in which the information set $\mathcal{S}_i^s = \mathcal{S}^s = (G_n, \mathbf{Z}_n)$ includes knowledge about everyone characteristics and the full network structure. Following the concept of Bayesian Nash equilibrium, we can define student i's equilibrium selection probability that maximizes the expected payoff given friends' average selection probabilities. Let $p_i^* = P(d_i = 1 \mid \mathcal{S}^s)$ be this equilibrium strategy for i and $\mathbf{p}_n^*$ be the corresponding vector.

$$
\begin{aligned}
p_i^* &= Pr\left(\mathbf{z}_i\gamma + \delta \mathbf{g}_i^* \mathbf{p}_n^* + v_i > 0\right) \qquad i = 1, \ldots, n \\
d_i &= 1\left(v_i > -\mathbf{z}_i\gamma - \delta \mathbf{g}_{i,n}^* \mathbf{p}_n^*\right).
\end{aligned}
\tag{2.2.2}
$$

For example, using the network in Fig. 2.1 again, we can describe students' uncertainty in making decisions to take a course in Science. Since they do not observe friends' actual choices, they instead guess about their friends' selection probabilities:

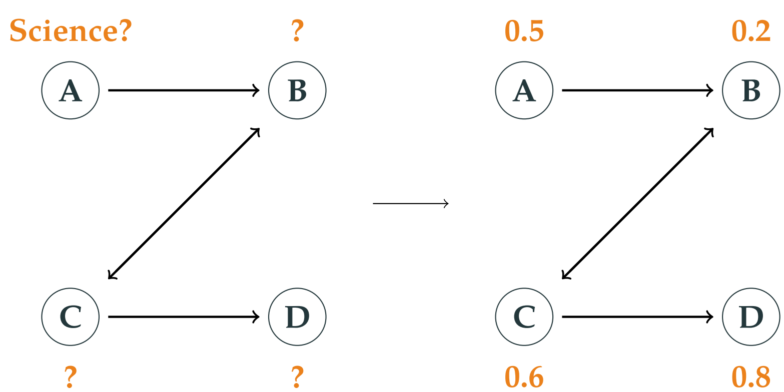


Figure 2.2: Illustration of first stage binary choice decision under incomplete information in a network

An important feature of this model is that we assume that v_i is idiosyncratic shock that is independent of the network structure as well as the covariates. Even though

[4]Even though i may call j to confirm whether j takes Social Studies this semester, we assume that there is a cost to extracting this extra information. The total cost to get confirmation of all 7 friends of i may exceed the simple guess about friends' selection.

equilibrium probabilities are dependent among students, the actual selection decisions are (conditionally) independent of each other.

The selection decision detailed in Eq. (2.2.2) has been studied by Xu (2018); Yang and Lee (2017). We refer to this model as a strategic binary choice model with incomplete information. Unique equilibrium conditions include a distributional assumption of the idiosyncratic shock, the limit on the strength of peer effects, as well as some conditions on the exogeneity and sparsity of network structure.

2.2.2 Linear outcome stage

Once students take subject courses, we can define another reduced-form utility function for the outcome stage:

$$U_i^o(\mathbf{y}_n \mid \mathbf{d}_n) = \left[\lambda \mathbf{g}_i^{o*}\mathbf{y}_n^* + \mathbf{x}_i\boldsymbol{\beta} + u_i\right] y_i^* - \frac{y_i^{*2}}{2} \tag{2.2.3}$$

then the reduced-form of the (potential) GPA level realized once students take the classes is expressed as:

$$\begin{aligned} y_i^* &= \lambda \mathbf{g}_i^{o*}\mathbf{y}_n^* + \mathbf{x}_i\boldsymbol{\beta} + u_i \\ \mathbf{y}_n^* &= \lambda G_n^{o*}\mathbf{y}_n^* + \mathbf{X}_n\boldsymbol{\beta} + \mathbf{u}_n \end{aligned} \tag{2.2.4}$$

We assume that there is a direct influence from the average friends' GPAs on one's own GPA as captured by parameter λ. If $\lambda > 0$, this can be interpreted that the increase in average of friends' GPAs benefits students' own performance. If $\lambda < 0$, on the other hand, it maybe the case that it is more difficult to achieve higher GPA.[5] To highlight the possible transformation of the peer structure once students select into courses in a subject, we denote the new adjacency matrix by G_n^o and the row-normalize version G_n^{o*}.

[5]In models that theorize about the nature of peer effects, Blume et al. (2015) specifies a similar utility function in which there is an additional social cost term so that the 'endogenous' peer effects are driven by conformity. Boucher and Fortin (2016) define a similar utility function to this model but restrict $\lambda > 0$ so that the spillover in network is due to complementarity. Fruehwirth (2014) assumes instead that there exist 'endogenous' peer effects in unobservables, such as effort, which in turns determine the outcome so that the final outcome observed by researchers is a proxy that exhibits dependence.

So far we have defined the potential GPA y_i^* for all students. However, for the students who do not take any course in the subject, their potential GPAs are observed neither by their friends in the network nor by researchers. Note that the unobservability nature of this outcome is different from that in the selection stage. In the first period, the actual selection is unobserved due to imperfect knowledge temporarily but will eventually be known as the semester starts. In the second period, the potential outcomes are not fully observed because some potential GPAs are not relevant for some students and will not be realized for the whole period. We argue that the peer effects come from some interaction between realized GPAs rather than potential GPAs during the semester. For example, it may be the case that since students take courses in the same subject, they do homework together and discuss on the same subject, which produces a synergy effect. With other friends who do not take any course in the subject that semester, students may devote time spent together to other common interests. As a result, we model such that the peer effects in subject-specific GPAs would occur only among students who actually take courses in the same subject at the same time.

Using the four-student example, suppose A, B, C all choose to take courses in Science. As a result, Science GPAs are realized for these three individuals only. No one observes D's potential outcomes. We can describe the new peer structure within the subject as:

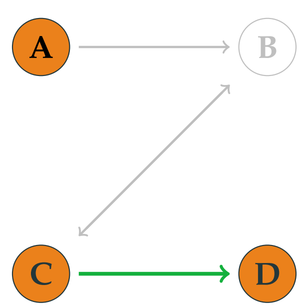


Figure 2.3: Example of a new network structure influenced by first stage choices of network members

This type of peer effects based on realized outcomes only can be captured by a new network structure G_n^o in which all the links to and from students who do not take any course in the subject is set to zero. In other words, $g_{ij} = d_i g_{ij} d_j$, and the for the normalized version $g_{ij}^* = d_i \dfrac{g_{ij} d_j}{\sum_j g_{ij} d_j}$ if $\sum_j g_{ij} d_j \neq 0$, and 0 otherwise. This new structure of peers that bases on the original friendship and is influenced by the selection decision in the first period.

The reduced form equation for GPAs is also called a Spatial Autoregressive model in the spatial literature. If peer effects is not too large, then if $S_n(\lambda) = (I_n - \lambda G_n^{o*})$ is invertible, then the system of equations $\mathbf{y}_n^*$ is stable and the (further) reduced form can also be written as:[6]

$$\mathbf{y}_n^* = S_n^{-1}(\lambda)\,[\mathbf{X}_n\boldsymbol{\beta} + \mathbf{u}_n] \tag{2.2.5}$$

Finally, we hypothesize that there is a relationship between selection decisions and the realized GPAs. For example, there is some common interest or motivation that influences both selection decision and academic performance. This can be translated to a positive correlation between two idiosyncratic shocks v_i and u_i in Eq. (2.2.2) and Eq. (2.2.4). In this sense, our model shares the same spirit with more traditional network formation models where endogeneity of the network is a result of some unobserved factor that influences both outcome and network. Here the unobserved factor influences selection decision that indirectly has consequences on the network structure.[7]

Our interest is the estimation of peer effects in the subject selection stage (δ) and in the subject-specific outcome stage (λ). The next section details the esimation.

2.3 Estimation method

Eq. (2.2.2) and Eq. (2.2.5) are the reduced form equations for estimation of our model. The parameters $\boldsymbol{\theta}^d = (\delta, \boldsymbol{\gamma}')'$ in the selection stage of the strategic binary choice with incomplete information can be estimated by Maximum Likelihood with nested fixed point

[6] Appendix **A.1** provides details for the incomplete information setting.

[7] The separation of two utility functions in the selection stage and outcome stage means that it is not the case that the only reason that students select courses in a subject is to maximize grades within that subject. In other words, our model is not a model of strategic network evolution where agents consciously construct a new network structure based on some underlying interest under constraint (no new links). For such type of models, we need to write a composite utility function $U_i(\mathbf{y}_n(\mathbf{d}_n))$ in which individuals should be aware of how peer structure and corresponding GPA outcomes will change in the second period. The optimization requires taking into account all possible realizations of peer structure. The utility function is likely to be highly nonlinear and whether there exists equilibria is an open question and interesting to explore in future research.

as suggested by Yang and Lee (2017).

The parameters in the outcome stage $\boldsymbol{\theta}^y$ can be estimated by nonlinear least squares. However, we need to give some attention to Eq. (2.2.5), as the new network G_n^{o*} is a function of all selection decisions of all students in the network in the previous period. Specifically, we need to control for the bias:

$$E\left[S_n^{-1}(\lambda)\mathbf{u}_n \mid \mathbf{d}_n, G_n\right] = S_n^{-1}(\lambda)E[\mathbf{u}_n \mid \mathbf{d}_n, G_n] \tag{2.3.1}$$

Using control function approach, we add and substract the bias term in Eq. (2.3.1) from the equation (2.2.5):

$$\mathbf{y}_n^* = S_n^{-1}(\lambda)\left[\mathbf{X}_n\boldsymbol{\beta} + E[\mathbf{u}_n \mid \mathbf{d}_n, G_n] + \underbrace{\mathbf{u}_n - E[\mathbf{u}_n \mid \mathbf{d}_n, G_n]}_{\boldsymbol{\epsilon}_n}\right] \tag{2.3.2}$$

Due to the incomplete information setting in the selection stage, we have each student's selection decision is conditionally independent given original network structure and control variables. As a result, $E(u_i \mid \mathbf{d}_n) = E(u_i \mid d_i)$. With some additional assumption on the relationship between v_i and u_i, we can construct terms similar to Heckman's inverse Mills ratio.

Note that $E\left[\boldsymbol{\epsilon}_n \mid \mathbf{d}_n, G_n\right] = 0$ by construction. Denote the number of unrealized outcomes ($\mathbf{y}_n^0$) and realized outcomes ($\mathbf{y}_n^1$) of length n^0 and n^1 respectively. Let J_n^1 be the $(n_1 \times n)$ selection matrix that extract (observed) realized outcomes $\mathbf{y}_n^1$ from $\mathbf{y}_n^*$ so that $\mathbf{y}_n^1 = \mathbf{y}_n^{1*} = J_n^1\mathbf{y}_n^*$.[8] The estimating equation is:

$$J_n^1\mathbf{y}_n^* = J_n^1 S_n^{-1}(\lambda)\left[\mathbf{X}_n\boldsymbol{\beta} + E(\mathbf{u}_n \mid \mathbf{d}_n, G_n)\right] + J_n^1 S_n^{-1}(\lambda)\boldsymbol{\epsilon}_n \tag{2.3.3}$$

with $E(J_n^1 S_n^{-1}(\lambda)\boldsymbol{\epsilon}_n) = 0$ since J_n^1 is a function of $\mathbf{d}_n$ and the exogenous G_n. As we correct for the bias in $\mathbf{u}_n$ conditional on $\mathbf{d}_n$, we solve two problems at the same time: endogeneity of network structure and selection, as both processes are related to just one type of selection

[8] This can be done by constructing an identity matrix I_n and removing the row indices corresponding to individuals having $d_i = 0$.

captured in $\mathbf{d}_n$.[9]

The bias correction in our model differs from that in other works on sample selection in the spatial literature which correct for the bias from $E(u_i \mid d_i = 1)$. For example, Rabovic and Cizek (2016) motivate an example similar to the following scenario, which also falls under the education setting. Suppose all students in a school are required to take the same test. Unfortunately, not all students show up on the exam day. Students may be influenced by friends when they decide whether to take the test. Nevertheless, while the academic performance proxied by the test scores of these no-show students are missing to the researchers, the performance may be known to their friends. Even though these students do not take exam, their performance may still affect and be affected by other friends who do take the exam. As a result, the network structure is likely to stay constant. The exogeneity of network in the outcome stage implies that the outcomes are missing to researchers only, while network members may still (directly or indirectly) observe it for everyone. Estimation of these models turns out to be a lot more intensive. For example, now we need to correct for each observed individual i: $\left[S_n^{-1}(\lambda)\right]_{ii} E(u_i \mid d_i = 1)$. Computationally, one needs to store the (dense) matrix $S_n^{-1}(\lambda)$ and extract diagonal elements of it, which is not feasible for large sample. Alternatively, we can do n times the approximation of $\mathbf{e}_i' S_n^{-1}(\lambda)\mathbf{e}_i E(u_i \mid d_i = 1)$ for $\mathbf{e}_i$ is the ith unit vector of the standard basis. The faster optimization in our setting is due to the fact that all error terms have the same conditioning set.

The main technical difficulty that our empirical model generates is that the peer structure in the outcome stage cannot be non-stochastic, as is often assumed in the literature. However, by working with the asymptotic where the sample size grows implies the increase in the number of independent subnetworks, instead of the setting in which there is a single large network, we can resolve the matter and show consistency of the two step estimator. The key assumption about network structure that makes the estimator valid is the following:

Assumption 1 (Network structure and asymptotic) *The sample consists of many indepen-*

[9] Technically, we can specify another binary variable h_i that further describes a missing data problem among the realized outcomes.

dent subnetworks, so that any observation from subnetwork k does not have any link, direct or indirect to another observation from subnetwork l. The size of each subnetwork is bounded by $c_g < \infty$ *as* $n \to \infty$. *Further more, network structure* G_n *is strictly exogenous.*

Assumption **1** imposes some sparsity to the network structure. The bounded size of each subnetwork implies that G_n is absolutely bounded in row and column sum.[10] The row boundedness of G_n means that an individuals's out-degree does not grow to infinity, i.e. each individual has a limited number of people he/she considers as friends even when the overall network size increases. The column boundedness of G_n further limits an individual's in-degree, i.e. each individual cannot be considered as a friend by 'too' many other people in the network as sample size goes to infinity. In the world wide web and modern online social network such as Twitter or Instagram this assumption is often violated, as some important members can have a very large number of followers. For our setting of high school student networks, there is a clear physical and institutional boundary that these two boundedness conditions are satisfied.[11]

The full list of the remaining conditions are provided below. These include some regularity conditions on covariates and the distributions of error terms, as well as the possible values of parameters that the model can admit. One special note is that since we have sample selection, we need to have a non-diminishing selection rate so that when the sample size is increased, there is actually new information. Naturally, non-perfect selection is also required to identify the parameters in the selection stage.

Assumption 2 *Regularity conditions: covariates and error terms.*

1. *The elements of* $\mathbf{Z}_n$ ($\mathbf{X}_n$) *are absolutely uniformly bounded in n;* $\mathbf{Z}_n$ ($\mathbf{X}_n$) *has full rank and* $\lim_{n\to\infty} \frac{1}{n}\mathbf{Z}_n'\mathbf{Z}_n$ ($\lim_{n\to\infty} \frac{1}{n}\mathbf{X}_n'\mathbf{X}_n$) *exists and is non-singular.*

[10] A sequence of square matrices $\{A_n\}$ where $A_n = [a_{ij}]$ is said to be uniformly bounded in row (column) sums in absolute value if the sequene of row sum matrix norm $||A_n||_\infty = \max_i \sum_{j=1}^{n} |a_{ij}|$ (column sum matrix norm $||A_n||_1 = \max_j \sum_{i=1}^{n} |a_{ij}|$) is bounded (Horn and Johnson, 1985).

[11] For exposition, so far we have assumed 'endogenous' peer effects only. We can extend to include the contextual peer effects such as by replacing $\mathbf{X}_n$ by $\widetilde{\mathbf{X}}_n = [\iota_n, \mathbf{X}_n, G_n^*\mathbf{X}_n]$ for ι_n be a column vector of ones of length n. The network structure for contextual peer effects does not have to change in the same way the network structure for 'endogenous' peer effects does. Since the contextual effects are described by a different peer structure, the model does not suffer from the reflection problem as described by Manski (1993). Practically, however, these two structures may still be highly correlated.

2. $\lim_{n\to\infty} \frac{n^1}{n} = \frac{\sum_i d_i}{n} = \alpha$ *for* α *be a finite positive constant.*

3. v_i *are iid distributed as standard normal distribution*

4. u_i *are iid such that* $E(u_i \mid v_i) = \rho v_i$.

5. v_i, u_i *are independent of* $\mathbf{X}_n, Z_n, G_n$.

Assumption 3 *Regularity conditions: parameters*

1. *The parameter space* Θ *is compact.*

2. *Parameter space of* δ, $\Delta \subset (-\sqrt{2\pi}, \sqrt{2\pi})$.

3. *Parameter space of of* λ, Λ *is such that* $\sup_{\lambda\in\Lambda} |\lambda| \sup_n ||G_n^{o*}||_\infty < 1$.

4. $S_n^{-1}(\lambda)$ *uniformly bounded in row and sum column in* $\lambda \in \Lambda$ *as* $n \to \infty$

In our model, the row normalized matrices G_n^* and G_n^{o*} are used in the reduced form outcomes in both stages. Effectively, the row normalization operation imposes uniformly bounded in row sum of G_n^*, G_n^{o*}. We have the number of (effective) friends in the second stage is also bounded by c_G since $\mathcal{G}_i^o \subseteq \mathcal{G}_i$. As G_n is bounded in column sum, we will have G_n^{o*} is also bounded in column sum, since D_n is a diagonal matrix. The column sum of jth row of G_n^o is, for $G_n^o = G_n D_n$

$$\sum_i g_{ij}^o = \sum_i \sum_k g_{ik} d_{kj} = \sum_i g_{ij} d_{jj} \leq \sum_i g_{ij}$$

The column sum of jth row of G_n^{o*} is, since Ω_n is diagonal matrix with elements non-negative, at most 1, for $G_n^{o*} = \Omega_n G_n D_n$

$$\sum_i g_{ij}^{o*} = \sum_i \sum_k \omega_{ik} g_{kj}^o \leq \sum_i \omega_{ii} g_{ij}^o \leq \sum_i g_{ij}^o \leq \sum_i g_{ij}$$

Let $\mathcal{W}_i^* = \sum_{j\neq i} |w_{ij}^*| = \sum_{j\neq i} w_{ij}^*$. By construction, $\mathcal{W}_i^* \in \{0,1\}$. We have Gershgorin disc $D(w_{ii}^*, \mathcal{W}_i^*) \in \mathbb{C}$ be a closed disc centered at w_{ii}^* with radius $\mathcal{W}_i^*$. By Gershgorin circle theorem, every eigenvalue of G_n^{o*} lies within at least one of the Gershgorin discs $D(0, \mathcal{W}_i^*)$ since $w_{ii}^* = 0$ for all $i \in n$. [12]

The maximum eigenvalue of G_n^{o*} is 1, a sufficient condition for $S_n(\lambda)$ to be invertible is such that the maximum magnitude of eigenvalue of $\lambda_0 G_n^{o*}$ is smaller than 1. Therefore, we can restrict $|\lambda_0| < 1$, i.e. $\Lambda \subset (-1, 1)$. Assumption 3(1) is required for the equilibrium of the linear second stage.

Assumption 4 *Identification conditions*

1. $E\left[(\mathbf{z}_i, \mathbf{g}_i^* \mathbf{p}_n^*)'(\mathbf{z}_i, \mathbf{g}_i^* \mathbf{p}_n^*)\right]$ *has full rank for all n sufficiently large.*

2. *Let* $\mathbf{m}_n = J_n^1 S_n^{-1}(\lambda)[\mathbf{X}_n(\boldsymbol{\theta}_0^d), G_n^{o*} S_n^{-1}(\lambda_0)\mathbf{X}_n(\boldsymbol{\theta}_0^d)\beta_0]$ *where* $\mathbf{X}_n(\boldsymbol{\theta}_0^d) = [\mathbf{X}_n, \mathbf{b}_n(\boldsymbol{\theta}_0^d)]$. *Full column rank of* $\lim_{n\to\infty} \frac{1}{n} E(\mathbf{m}_n' \mathbf{m}_n)$ *for any value* λ *in* Λ.

Under assumption 1, assumption 2(1), (3), (5), assumption 3(1), (2), and assumption 4(1), we can verify the existence of a unique equlibrium and identification of parameters of the binary model in the first stage as in Yang and Lee (2017).

Consider assumption 4(2): for identification, we require that the bias term should not be a linear function of covariates. As with the original Heckman (1979), while the non-linear functional form as well as $\mathbf{p}_n^*$ of the selection term helps, it may still be highly correlated with the rest of the covariates. It is recommended that we have some elements in $\mathbf{Z}_n$ but not in $\mathbf{X}_n$. The importance of having at least one instrumental variable for exclusion restriction will be demonstrated in the simulation analysis.[13]

[12] Note that Gershgorin circle theorem does not require the matrix to be symmetric.

[13] Suppose instead of $\mathbf{X}_n$, we assume $\widetilde{\mathbf{X}}_n = [\boldsymbol{\iota}_n, \mathbf{X}_n, G_n^* \mathbf{X}_n, \mathbf{b}_n(\gamma, \delta)]$. The corresponding parameters are $\alpha, \boldsymbol{\beta}^{own}, \boldsymbol{\beta}^{fr}, \rho$:

$$\mathbf{y}_n^* = S_n^{-1}(\lambda)\left[\alpha\boldsymbol{\iota}_n + \mathbf{X}_n\boldsymbol{\beta}^{own} + G_n^*\mathbf{X}_n\boldsymbol{\beta}^{fr} + \underbrace{E(\mathbf{u}_n \mid \mathbf{d}_n, \mathbf{Z}_n, G_n)}_{\rho\mathbf{b}_n}\right] + S_n^{-1}(\lambda)\boldsymbol{\epsilon}_n \tag{2.3.4}$$

$$= \alpha S_n^{-1}(\lambda)\boldsymbol{\iota}_n + \mathbf{X}_n\boldsymbol{\beta}^{own} + \lambda G_n^{o*} S_n^{-1}(\lambda)\mathbf{X}_n\boldsymbol{\beta}^{own} + S_n^{-1}(\lambda) G_n^* \mathbf{X}_n \boldsymbol{\beta}^{fr} + S_n^{-1}(\lambda)\rho\mathbf{b}_n + S_n^{-1}(\lambda)\boldsymbol{\epsilon}_n \tag{2.3.5}$$

Since $S_n^{-1}(\lambda) = (I_n - \lambda G_n^{o*})^{-1} = \sum_{k=0}^{\infty}(\lambda G_n^{o*})^k = I_n + \lambda G_n^{o*} S_n^{-1}(\lambda)$. When the peer structure for contextual effects is also G_n^{o*}, Bramoullé et al. (2009) show that a sufficient condition for identification of endogenous and contextual peer effects is that I_n, G_n^{o*}, G_n^{o*2} are linearly independent and that $\lambda\boldsymbol{\beta}^{own} + \boldsymbol{\beta}^{fr} \neq \mathbf{0}$, i.e. there is intransitive triads among the selected students, i.e. there is a non-mutual friend between two selected students that also select into the outcome stage. In this case, however, since $G_n^* \neq G_n^{o*}$ then

Theorem 1 (Consistency of outcome stage) *Given assumptions 1 to 4, $\boldsymbol{\theta}^y$ can be identified and estimated by non linear least square using a consitent estimate $\widehat{\boldsymbol{\theta}}^d$ of $\boldsymbol{\theta}^d$ in the first stage.*

The proof is shown in Appendix **A.3**.

With full distributional assumption of u_i, if researchers believe the outcome stage follows an incomplete information setting as in Appendix **A.1**, an alternative estimation method is the partial MLE.[14]

For inference, we use block bootstrap. In the spatial literature, residual bootstaps have been proposed (Jin and Lee, 2015). Some non-parametric bootstrapping methods for network data are developed for the case of a single large network (Bhattacharyya and Bickel, 2015; Thompson et al., 2016; Gel et al., 2017). In some settings, a special structure of data of multiple independent subnetworks of schools can be exploited so that we can use block bootstrap, which is a special case of the class of non-parametric bootstraps. In our estimation, we draw subnetworks with replacement for $B = 500$ times to estimate and form a 95% CI.

2.4 Monte Carlo simulation

In this section, we conduct several simulations to evaluate our estimator under different data generating processes (DGP), the performance of block bootstrap, and the usefulness of having the exclusion restriction assumption satisfied.

Instead of simulating network and covariates from scratch, which requires additional specifications of the network formation process and DGP of independent variables, we make use of a special data set, the National Longitudinal Study of Adolescent and Adult Health (Add Health) data. The details of this data source is discussed further in the second chapter. For now, it suffices to summarize that Add Health elicits network connection data as well as various demographic characteristics of high school students in 1994-1995 in multiple schools across the United States. In this simulation section, we collect network

$G_n^{o*} S_n^{-1}(\lambda) \neq S_n^{-1}(\lambda) G_n^*$, the separation of contextual and endogenous peer effects should be achieved more easily. However, since in our empirical application, selection rate is very high, it could be the case that G_n^* and G_n^{o*} are very similar.

[14]With incomplete information setting in the second stage, it is possible to write down the partial likelihood, which is based on (A.1.3).

links and two covariates: health (a rank from 1 to 5) and gender (binary). The variable gender serves as the exclusion restriction variable where it enters the selection stage but not the outcome stage. We consider subnetwork at school-grade level. This means that we construct network links among students attending the same grade in the same shool only. The total number of subnetworks available for analysis is around 200, which we view as the population of subnetworks.

We consider a growing (sample size) number of subnetworks: 20, 40, 80, and 160. For each sample size, we conduct 500 Monte Carlo simulations in which we draw subnetworks with replacement and simulate outcome variables according to the true data generating process (DGP.)

2.4.1 Model misspecifications

We considere four DGPs that differ in the nature of selection (random or endogenous) and in the fluidity of network structure (dynamic: only individuals selecting into the second stage (a subject such as Science) can influence the subject outcomes (Science GPAs) of others vs fixed: any individuals whether selecting or not selecting into the same subject (Science) can influence the outcomes of others).

1. DGP(i): The true model is selection and *dynamic* network, i.e., the model of the dissertation.

   In selection stage, the endogenous peer effects coefficient is set at $\delta = 0.2$. The constant for first stage is calibrated such that the average probability of selection without peer effects is 0.8.

   In outcome stage, the endogenous peer effects coefficient is $\lambda = 0.05$. Error terms u_i, v_i is distributed as bivariate normal with covariance 0.3, $\sigma_v = 1$ and $\sigma_u = 2$ so that $\rho = 0.6$.

   The rest of the coefficients in both stages are generated from $N(0, 1)$.

   Only individuals selecting into the outcome stage can influence each other, i.e. network is *dynamic*.

2. DGP(ii): The true model is random selection and dynamic network.

   In selection stage, every individual has the same probability 0.8 of selecting into the outcome stage.

   In outcome stage, the endogenous peer effects coefficient is $\lambda = 0.05$. u_i is distributed as normal with mean zero and $\sigma_u = 2$. The rest of the coefficients are generated from $N(0,1)$. Only individuals selecting into the outcome stage can influence each other, i.e. network is *dynamic*.

3. DGP(iii): The true model is random selection and fixed network.

   In selection stage, very individual has the same probability 0.8 of selecting into the outcome stage.

   In outcome stage, the endogenous peer effects coefficient is $\lambda = 0.05$. u_i is distributed as normal with mean zero and $\sigma_u = 2$. The rest of the coefficients are generated from $N(0,1)$. There are peer effects among selecting and non-selecting individuals, so that network is *fixed*. This is the case of missing dependent variable at random.

4. DGP(iv): The true model is endogenous selection and fixed network.

   In selection stage, the endogenous peer effects coefficient is set at $\delta = 0.2$. The constant for first stage is calibrated such that the average probability of selection without peer effects is 0.8.

   In outcome stage, the endogenous peer effects coefficient is $\lambda = 0.05$. error terms u_i, v_i is distributed as bivariate normal with covariance 0.3, $\sigma_v = 1$ and $\sigma_u = 2$ so that $\rho = 0.6$.

   The rest of the coefficients in both stages are generated from $N(0,1)$.

   There are peer effects among selecting and non-selecting individuals, so that network is *fixed*.

For each DGP, we estimate using three estimators, (I), (II), (III). (I) is our main two-step estimator where the first step is nested fixed point MLE for strategic binary choice with incomplete information and the second step is NLS that corrects for endogenous selection and with dynamic network. (I) is the correct estimator under DGP(i). (II) and (III) estimate the second stage only. (II) is NLS not correcting for selection and with dynamic network that is correct under DGP(ii), and (III) is NLS not correcting for selection and with fixed network that is correct under DGP(iii). Under DGP(iv), no estimator is correct.

For each sample size, we report the finite bias and RMSE for the endogenous peer effects λ, constant, covariate coefficient β, selection coefficient ρ. Since three estimators assume different network structure, instead of just focusing on the bias of coefficients, we also calculate other statistics, namely average total effects (ATEF) under true DGP and its estimates ($\widehat{\text{ATEF}}$) by these three different estimators.

With the endogenous peer effects, the interpretation of the marginal effects of covariates for SAR model will be different from the usual OLS model. We focus on the marginal effects of the second stage only. The reduced form of second stage of the model is:

$$\mathbf{y}_n = (I_n - \lambda G_n^{o*})^{-1}\mathbf{X}_n\boldsymbol{\beta} + (I_n - \lambda G_n^{o*})^{-1}\mathbf{u}_n$$

The marginal effect of $\mathbf{X}_k$ therefore is captured by the matrix $(I_n - \lambda G_n^{o*})^{-1}\beta_k$. Without peer effects, $(I_n - \lambda G_n^{o*})^{-1}$ is just the identity matrix so that the matrix of marginal effects is:

$$\begin{pmatrix} \beta^k & 0 & \cdots & 0 \\ 0 & \beta^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \beta^k \end{pmatrix}$$

which means the marginal impact of $x_{k,i}$ on y_i is homogenous across individuals and there is no spillover impacts of $x_{k,i}$ on y_j. However, since we have statistically significant peer effects λ, the marginal effects matrix will not be a diagonal matrix and can be dense with many non-zero elements, which can look like this

$$\begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1n} \\ s_{21} & s_{22} & \cdots & s_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ s_{n1} & s_{n2} & \cdots & s_{nn} \end{pmatrix} \beta_k$$

The overall impact of $x_{k,i}$ on everyone in the network is captured by the ith column. Therefore, marginal impacts of each exogenous variable will be heterogenous.

Reporting the full $n \times n$ matrix for each exogenous variable is burdensom and not very informative. Assuming selection decisions remain fixed, there are two types of marginal effects with respect to the outcome that are of interest: (1) Average of direct/own effects (average of diagonal elementstimes β_k); (2) Average of total effects (average of column sums times β_k), which is the average of effects on everyone in the network (including self).[15]

In this section, we focus on finite biases and MC standard deviations for ATEF (average total effects) which is the average of column sum of $(I_n - \lambda G_n^{o*})^{-1}$, and ATEFX which is the average of column sum of $(I_n - \lambda G_n^{o*})^{-1}\beta$. G_n^{o*} takes different network structure based on the assumption of dynamic or fixed network of estimators. Furthermore, we calculate these marginal effects by averaging over selecting individuals only since we want to see whether these estimators can capture the overall spillovers among the observed outcomes.

The simulation results are in Tables **2.1** to **2.4**. We highlighted the estimator that is correct under each DGP. The corresponding correct estimator performs well as expected.

Estimates from (II) and (III) are fairly similar throughout. This is clear when we set the selection rate to be 0.8, the dynamic network and its original version are highly correlated,

[15]We focus on the marginal effects with respect to outcome only, given the decision made in the first stage as fixed. Theoretically, analyzing the full marginal impacts of $\mathbf{x}_k$ in both stages entails two tasks. First, changing $\mathbf{x}_k$ will change the equilibrium probabilities in the selection stage, which takes a highly non-linear form. Such marginal effects have been calculated only for the case where error in the first stage is assumed to follow logit (Yang and Lee (2017)), but not under normal distribution as we consider in this paper. Secondly and more importantly, the change in equilibrium probabilities subsequently changes the actual decision and the corresponding new network structure in the second stage. The calculation of the (expected) marginal effects of $\mathbf{x}_k$ therefore would require taking the expectation of effects in the second stage with respect to the space of all possible realization of selection decision $\mathbf{d}_n$. Simulations may help.

but this is also the case when we reduce the average selection rate of the first stage to 0.2. The constant term is usually the term that suffers from misspecification in network structure.

Under DGP(ii) and DPG(iii) where selection is indeed random, regardless of the exact network structure, (I) introduces large biases for the constant and selection term with high standard deviations. On the other hand, the marginal effects as captured by ATEFX are not too far off from the estimates of correct estimators. Most imterestingly, under DGP(iv) where there is endogenous selection but the network in the outcome stage remains fixed, (I) outperforms the other two for all coefficients estimates as well as marginal effects. This benefit comes mainly from the low bias in coefficient for covariates. In Table **2.5**, we keep selection rate at 0.8 but increases endogenous peer effects to $\lambda = 0.5$ for DGP(iv), the story remains the same, as the gain from low bias in β helps characterizing the marginal effects much better than estimators that do not control for selection.

Overall, it seems that SAR model is more sensitive to endogeneity than to the specification of the network matrix.

Table 2.1: Simulation results for estimation under DGP(i): Endogenous selection, dynamic network

			ng = 20		ng = 40		ng = 80		ng = 160	
			Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Selection rate: 0.8										
(I)	λ	0.050	-6e-04	[3e-04]	-3e-04	[2e-04]	-2e-04	[8e-05]	-2e-04	[4e-05]
	Cons	-0.423	-0.009	[3e-02]	-0.004	[1e-02]	0.001	[6e-03]	0.004	[4e-03]
	β	-1.774	0.001	[3e-03]	0.001	[1e-03]	0.000	[5e-04]	-0.002	[3e-04]
	ρ	0.600	0.011	[2e-02]	0.001	[1e-02]	-0.004	[5e-03]	-0.003	[2e-03]
	ATEF	1.040	-3e-04	[2e-04]	-2e-04	[1e-04]	-1e-04	[6e-05]	-2e-04	[3e-05]
	ATEFX	-1.845	1e-03	[4e-03]	2e-03	[2e-03]	3e-04	[8e-04]	-1e-03	[5e-04]
(II)	λ	0.050	1e-02	[5e-04]	1e-02	[3e-04]	1e-02	[3e-04]	1e-02	[2e-04]
	Cons	-0.423	0.476	[2e-01]	0.473	[2e-01]	0.474	[2e-01]	0.479	[2e-01]
	β	-1.774	-0.109	[1e-02]	-0.107	[1e-02]	-0.107	[1e-02]	-0.109	[1e-02]
	ρ	0.600	-	-	-	-	-	-	-	-
	ATEF	1.040	1e-02	[4e-04]	1e-02	[3e-04]	1e-02	[2e-04]	1e-02	[2e-04]
	ATEFX	-1.845	-1e-01	[2e-02]	-1e-01	[2e-02]	-1e-01	[2e-02]	-1e-01	[2e-02]
(III)	λ	0.050	2e-02	[6e-04]	2e-02	[5e-04]	2e-02	[3e-04]	2e-02	[3e-04]
	Cons	-0.423	0.481	[2e-01]	0.478	[2e-01]	0.477	[2e-01]	0.481	[2e-01]
	β	-1.774	-0.109	[1e-02]	-0.107	[1e-02]	-0.107	[1e-02]	-0.109	[1e-02]
	ρ	0.600	-	-	-	-	-	-	-	-
	ATEF	1.040	4e-03	[2e-04]	4e-03	[1e-04]	4e-03	[6e-05]	4e-03	[4e-05]
	ATEFX	-1.845	-1e-01	[2e-02]	-1e-01	[2e-02]	-1e-01	[1e-02]	-1e-01	[1e-02]
Selection rate: 0.2										
(I)	λ	0.050	-2e-03	[6e-04]	-7e-04	[3e-04]	-7e-05	[1e-04]	-5e-05	[6e-05]
	Cons	-0.423	-0.031	[1e-01]	-0.011	[7e-02]	-0.018	[3e-02]	0.007	[2e-02]
	β	-1.774	0.003	[8e-03]	0.002	[4e-03]	0.004	[2e-03]	-0.002	[1e-03]
	ρ	0.600	0.015	[3e-02]	0.005	[1e-02]	0.011	[8e-03]	0.000	[4e-03]
	ATEF	1.031	-9e-04	[2e-04]	-4e-04	[1e-04]	1e-05	[5e-05]	-2e-06	[3e-05]
	ATEFX	-1.829	4e-03	[1e-02]	3e-03	[5e-03]	4e-03	[2e-03]	-2e-03	[1e-03]
(II)	λ	0.050	1e-02	[7e-04]	1e-02	[5e-04]	2e-02	[3e-04]	1e-02	[3e-04]
	Cons	-0.423	1.029	[1e+00]	1.032	[1e+00]	1.031	[1e+00]	1.040	[1e+00]
	β	-1.774	-0.197	[4e-02]	-0.194	[4e-02]	-0.193	[4e-02]	-0.196	[4e-02]
	ρ	0.600	-	-	-	-	-	-	-	-
	ATEF	1.031	9e-03	[3e-04]	1e-02	[2e-04]	1e-02	[1e-04]	1e-02	[1e-04]
	ATEFX	-1.829	-2e-01	[5e-02]	-2e-01	[5e-02]	-2e-01	[5e-02]	-2e-01	[5e-02]
(III)	λ	0.050	1e-02	[1e-03]	2e-02	[7e-04]	2e-02	[5e-04]	1e-02	[3e-04]
	Cons	-0.423	1.041	[1e+00]	1.046	[1e+00]	1.044	[1e+00]	1.050	[1e+00]
	β	-1.774	-0.193	[4e-02]	-0.191	[4e-02]	-0.190	[4e-02]	-0.193	[4e-02]
	ρ	0.600	-	-	-	-	-	-	-	-
	ATEF	1.031	-1e-02	[2e-04]	-1e-02	[2e-04]	-1e-02	[1e-04]	-1e-02	[1e-04]
	ATEFX	-1.829	-2e-01	[4e-02]	-2e-01	[3e-02]	-2e-01	[3e-02]	-2e-01	[3e-02]

(I) Bias correction with dynamic network. (II) No bias correction with dynamic network. (III) No bias correction with fixed network. 500 Monte Carlo simulations.

Table 2.2: Simulation results for estimation under DGP(ii): Random selection, dynamic network

			ng = 20		ng = 40		ng = 80		ng = 160	
			Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Selection rate: 0.8										
(I)	λ	0.050	-6e-04	[5e-04]	2e-04	[3e-04]	4e-04	[1e-04]	-2e-04	[5e-05]
	Cons	-0.423	0.048	[6e+00]	-0.160	[1e+01]	0.097	[6e+00]	0.008	[9e+00]
	β	-1.774	0.002	[9e-03]	0.002	[5e-03]	0.001	[1e-03]	-0.000	[9e-04]
	ρ	0.000	-0.173	[5e+01]	0.442	[8e+01]	-0.269	[4e+01]	-0.019	[8e+01]
	ATEF	1.040	-1e-04	[4e-04]	3e-04	[2e-04]	4e-04	[9e-05]	-1e-04	[4e-05]
	ATEFX	-1.844	2e-03	[1e-02]	1e-03	[6e-03]	-2e-05	[2e-03]	-2e-04	[1e-03]
(II)	λ	0.050	-3e-04	[3e-04]	5e-04	[2e-04]	-1e-04	[8e-05]	-1e-04	[4e-05]
	Cons	-0.423	-0.007	[1e-02]	-0.003	[6e-03]	-0.002	[3e-03]	0.001	[1e-03]
	β	-1.774	0.001	[1e-03]	0.001	[8e-04]	0.001	[4e-04]	-0.000	[2e-04]
	ρ	0.000	-	-	-	-	-	-	-	-
	ATEF	1.040	-4e-05	[2e-04]	5e-04	[1e-04]	-7e-05	[5e-05]	-6e-05	[3e-05]
	ATEFX	-1.844	1e-03	[2e-03]	6e-04	[1e-03]	9e-04	[6e-04]	-2e-04	[3e-04]
(III)	λ	0.050	-3e-03	[3e-04]	-3e-03	[2e-04]	-3e-03	[9e-05]	-3e-03	[5e-05]
	Cons	-0.423	-0.010	[1e-02]	-0.009	[6e-03]	-0.006	[3e-03]	-0.004	[2e-03]
	β	-1.774	0.002	[1e-03]	0.002	[8e-04]	0.002	[4e-04]	0.001	[2e-04]
	ρ	0.000	-	-	-	-	-	-	-	-
	ATEF	1.040	-8e-03	[2e-04]	-8e-03	[2e-04]	-9e-03	[1e-04]	-9e-03	[1e-04]
	ATEFX	-1.844	2e-02	[2e-03]	2e-02	[1e-03]	2e-02	[8e-04]	2e-02	[5e-04]
Selection rate: 0.2										
(I)	λ	0.050	-4e-04	[1e-03]	-3e-05	[6e-04]	-9e-04	[3e-04]	2e-04	[2e-04]
	Cons	-0.423	-0.550	[1e+02]	0.288	[3e+02]	-0.315	[1e+02]	-0.835	[3e+02]
	β	-1.774	0.004	[2e-02]	-0.006	[7e-02]	-0.004	[9e-03]	-0.003	[4e-03]
	ρ	0.000	0.382	[6e+01]	-0.196	[2e+02]	0.228	[7e+01]	0.601	[2e+02]
	ATEF	1.021	9e-05	[2e-04]	1e-04	[1e-04]	-3e-04	[5e-05]	1e-04	[3e-05]
	ATEFX	-1.811	4e-03	[2e-02]	-7e-03	[7e-02]	-3e-03	[1e-02]	-4e-03	[5e-03]
(II)	λ	0.050	6e-05	[1e-03]	3e-05	[5e-04]	-9e-04	[2e-04]	2e-05	[1e-04]
	Cons	-0.423	-0.014	[4e-02]	-0.018	[2e-02]	-0.003	[1e-02]	0.004	[5e-03]
	β	-1.774	0.004	[6e-03]	0.008	[4e-03]	0.002	[2e-03]	-0.002	[8e-04]
	ρ	0.000	-	-	-	-	-	-	-	-
	ATEF	1.021	3e-04	[2e-04]	1e-04	[1e-04]	-3e-04	[5e-05]	5e-05	[3e-05]
	ATEFX	-1.811	3e-03	[7e-03]	8e-03	[4e-03]	2e-03	[2e-03]	-2e-03	[9e-04]
(III)	λ	0.050	-3e-02	[2e-03]	-3e-02	[1e-03]	-3e-02	[1e-03]	-3e-02	[9e-04]
	Cons	-0.423	-0.027	[5e-02]	-0.035	[3e-02]	-0.018	[1e-02]	-0.011	[6e-03]
	β	-1.774	0.008	[6e-03]	0.011	[4e-03]	0.005	[2e-03]	0.001	[8e-04]
	ρ	0.000	-	-	-	-	-	-	-	-
	ATEF	1.021	-2e-02	[3e-04]	-2e-02	[3e-04]	-2e-02	[3e-04]	-2e-02	[3e-04]
	ATEFX	-1.811	4e-02	[8e-03]	4e-02	[5e-03]	4e-02	[3e-03]	3e-02	[2e-03]

(I) Bias correction with dynamic network. (II) No bias correction with dynamic network. (III) No bias correction with fixed network. 500 Monte Carlo simulations.

Table 2.3: Simulation results for estimation under DGP(iii): Random selection, fixed network

			ng = 20		ng = 40		ng = 80		ng = 160	
			Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Selection rate: 0.8										
(I)	λ	0.050	-1e-02	[8e-04]	-1e-02	[5e-04]	-1e-02	[3e-04]	-1e-02	[3e-04]
	Cons	-0.423	0.099	[8e+00]	-0.233	[2e+01]	-0.043	[2e+01]	0.159	[5e+01]
	β	-1.774	0.001	[8e-03]	-0.001	[8e-03]	-0.002	[3e-03]	0.001	[5e-03]
	ρ	0.000	-0.450	[6e+01]	0.525	[1e+02]	0.010	[1e+02]	-0.591	[4e+02]
	ATEF	1.033	-4e-03	[4e-04]	-4e-03	[2e-04]	-4e-03	[1e-04]	-4e-03	[7e-05]
	ATEFX	-1.833	8e-03	[1e-02]	5e-03	[9e-03]	4e-03	[3e-03]	8e-03	[5e-03]
(II)	λ	0.050	-7e-03	[3e-04]	-6e-03	[2e-04]	-7e-03	[1e-04]	-7e-03	[9e-05]
	Cons	-0.423	-0.033	[1e-02]	-0.028	[6e-03]	-0.027	[4e-03]	-0.025	[2e-03]
	β	-1.774	-0.001	[1e-03]	-0.000	[8e-04]	-0.001	[4e-04]	-0.002	[2e-04]
	ρ	0.000	-	-	-	-	-	-	-	-
	ATEF	1.033	9e-04	[2e-04]	1e-03	[1e-04]	9e-04	[5e-05]	8e-04	[3e-05]
	ATEFX	-1.833	-2e-03	[2e-03]	-3e-03	[1e-03]	-3e-03	[5e-04]	-4e-03	[3e-04]
(III)	λ	0.050	3e-05	[3e-04]	-6e-05	[2e-04]	-2e-04	[8e-05]	-3e-04	[4e-05]
	Cons	-0.423	-0.006	[1e-02]	-0.004	[6e-03]	-0.002	[3e-03]	0.000	[2e-03]
	β	-1.774	0.001	[1e-03]	0.001	[8e-04]	0.001	[4e-04]	-0.000	[2e-04]
	ρ	0.000	-	-	-	-	-	-	-	-
	ATEF	1.033	2e-04	[2e-04]	8e-05	[9e-05]	-7e-05	[4e-05]	-2e-04	[2e-05]
	ATEFX	-1.833	7e-04	[2e-03]	1e-03	[1e-03]	8e-04	[5e-04]	-5e-05	[3e-04]
Selection rate: 0.2										
(I)	λ	0.050	-4e-02	[3e-03]	-4e-02	[2e-03]	-4e-02	[2e-03]	-4e-02	[2e-03]
	Cons	-0.423	-0.688	[2e+02]	-0.120	[3e+02]	1.252	[4e+02]	0.582	[9e+02]
	β	-1.774	0.002	[3e-02]	-0.006	[5e-02]	-0.006	[3e-02]	-0.005	[2e-02]
	ρ	0.000	0.380	[9e+01]	-0.015	[2e+02]	-0.997	[2e+02]	-0.500	[4e+02]
	ATEF	1.008	-4e-03	[2e-04]	-3e-03	[1e-04]	-4e-03	[6e-05]	-4e-03	[4e-05]
	ATEFX	-1.789	8e-03	[3e-02]	7e-04	[5e-02]	1e-03	[3e-02]	1e-03	[2e-02]
(II)	λ	0.050	-3e-02	[2e-03]	-3e-02	[2e-03]	-3e-02	[1e-03]	-3e-02	[1e-03]
	Cons	-0.423	-0.138	[5e-02]	-0.141	[4e-02]	-0.126	[3e-02]	-0.119	[2e-02]
	β	-1.774	-0.002	[6e-03]	0.002	[3e-03]	-0.004	[2e-03]	-0.008	[8e-04]
	ρ	0.000	-	-	-	-	-	-	-	-
	ATEF	1.008	-4e-04	[2e-04]	-5e-04	[9e-05]	-9e-04	[4e-05]	-6e-04	[2e-05]
	ATEFX	-1.789	-1e-03	[7e-03]	3e-03	[4e-03]	-2e-03	[2e-03]	-7e-03	[9e-04]
(III)	λ	0.050	9e-04	[1e-03]	-7e-04	[6e-04]	-4e-04	[3e-04]	-3e-04	[2e-04]
	Cons	-0.423	-0.012	[4e-02]	-0.020	[2e-02]	-0.003	[1e-02]	0.004	[6e-03]
	β	-1.774	0.004	[6e-03]	0.008	[3e-03]	0.002	[2e-03]	-0.002	[8e-04]
	ρ	0.000	-	-	-	-	-	-	-	-
	ATEF	1.008	5e-04	[5e-05]	4e-05	[2e-05]	2e-05	[1e-05]	-2e-05	[5e-06]
	ATEFX	-1.789	3e-03	[6e-03]	8e-03	[4e-03]	2e-03	[2e-03]	-2e-03	[8e-04]

(I) Bias correction with dynamic network. (II) No bias correction with dynamic network. (III) No bias correction with fixed network. 500 Monte Carlo simulations.

Table 2.4: Simulation results for estimation under DGP(iv): Endogenous selection, fixed network

			ng = 20		ng = 40		ng = 80		ng = 160	
			Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Selection rate: 0.8										
(I)	λ	0.050	-9e-03	[4e-04]	-8e-03	[2e-04]	-8e-03	[2e-04]	-8e-03	[1e-04]
	Cons	-0.423	-0.040	[3e-02]	-0.035	[1e-02]	-0.030	[7e-03]	-0.027	[4e-03]
	β	-1.774	0.000	[3e-03]	0.001	[1e-03]	-0.000	[5e-04]	-0.002	[3e-04]
	ρ	0.600	0.010	[2e-02]	-0.000	[9e-03]	-0.005	[5e-03]	-0.003	[2e-03]
	ATEF	1.033	3e-04	[2e-04]	4e-04	[1e-04]	5e-04	[6e-05]	4e-04	[3e-05]
	ATEFX	-1.832	-2e-04	[4e-03]	2e-05	[2e-03]	-1e-03	[8e-04]	-3e-03	[5e-04]
(II)	λ	0.050	5e-03	[3e-04]	6e-03	[2e-04]	6e-03	[1e-04]	6e-03	[7e-05]
	Cons	-0.423	0.445	[2e-01]	0.442	[2e-01]	0.443	[2e-01]	0.448	[2e-01]
	β	-1.774	-0.110	[1e-02]	-0.108	[1e-02]	-0.108	[1e-02]	-0.110	[1e-02]
	ρ	0.600	-	-	-	-	-	-	-	-
	ATEF	1.033	1e-02	[4e-04]	1e-02	[3e-04]	1e-02	[2e-04]	1e-02	[2e-04]
	ATEFX	-1.832	-1e-01	[2e-02]	-1e-01	[2e-02]	-1e-01	[2e-02]	-1e-01	[2e-02]
(III)	λ	0.050	2e-02	[6e-04]	2e-02	[4e-04]	2e-02	[3e-04]	2e-02	[3e-04]
	Cons	-0.423	0.474	[2e-01]	0.471	[2e-01]	0.470	[2e-01]	0.474	[2e-01]
	β	-1.774	-0.109	[1e-02]	-0.106	[1e-02]	-0.107	[1e-02]	-0.109	[1e-02]
	ρ	0.600	-	-	-	-	-	-	-	-
	ATEF	1.033	1e-02	[3e-04]	1e-02	[2e-04]	1e-02	[2e-04]	1e-02	[1e-04]
	ATEFX	-1.832	-1e-01	[2e-02]	-1e-01	[2e-02]	-1e-01	[2e-02]	-1e-01	[2e-02]
Selection rate: 0.2										
(I)	λ	0.050	-3e-02	[1e-03]	-3e-02	[1e-03]	-3e-02	[8e-04]	-3e-02	[7e-04]
	Cons	-0.423	-0.164	[2e-01]	-0.145	[9e-02]	-0.152	[6e-02]	-0.126	[3e-02]
	β	-1.774	0.007	[8e-03]	0.006	[4e-03]	0.008	[2e-03]	0.002	[1e-03]
	ρ	0.600	0.028	[3e-02]	0.018	[2e-02]	0.024	[9e-03]	0.013	[4e-03]
	ATEF	1.016	-2e-03	[2e-04]	-2e-03	[1e-04]	-1e-03	[5e-05]	-1e-03	[3e-05]
	ATEFX	-1.802	1e-02	[1e-02]	1e-02	[5e-03]	1e-02	[2e-03]	4e-03	[1e-03]
(II)	λ	0.050	-1e-02	[7e-04]	-1e-02	[4e-04]	-1e-02	[2e-04]	-1e-02	[2e-04]
	Cons	-0.423	0.920	[9e-01]	0.922	[9e-01]	0.921	[9e-01]	0.930	[9e-01]
	β	-1.774	-0.198	[4e-02]	-0.195	[4e-02]	-0.194	[4e-02]	-0.197	[4e-02]
	ρ	0.600	-	-	-	-	-	-	-	-
	ATEF	1.016	7e-03	[3e-04]	8e-03	[2e-04]	8e-03	[1e-04]	8e-03	[9e-05]
	ATEFX	-1.802	-2e-01	[5e-02]	-2e-01	[5e-02]	-2e-01	[5e-02]	-2e-01	[5e-02]
(III)	λ	0.050	2e-02	[1e-03]	2e-02	[9e-04]	2e-02	[7e-04]	2e-02	[5e-04]
	Cons	-0.423	1.020	[1e+00]	1.024	[1e+00]	1.022	[1e+00]	1.029	[1e+00]
	β	-1.774	-0.194	[4e-02]	-0.191	[4e-02]	-0.190	[4e-02]	-0.193	[4e-02]
	ρ	0.600	-	-	-	-	-	-	-	-
	ATEF	1.016	7e-03	[2e-04]	7e-03	[1e-04]	7e-03	[8e-05]	7e-03	[6e-05]
	ATEFX	-1.802	-2e-01	[5e-02]	-2e-01	[5e-02]	-2e-01	[4e-02]	-2e-01	[4e-02]

(I) Bias correction with dynamic network. (II) No bias correction with dynamic network. (III) No bias correction with fixed network. 500 Monte Carlo simulations.

Table 2.5: Simulation results for estimation under DGP(iv): Endogenous selection, fixed network with $\lambda = 0.5$

			ng = 20		ng = 40		ng = 80		ng = 160	
			Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
(I)	λ	0.500	-7e-02	[6e-03]	-7e-02	[5e-03]	-7e-02	[5e-03]	-7e-02	[5e-03]
	Cons	-0.423	-0.449	[2e-01]	-0.446	[2e-01]	-0.432	[2e-01]	-0.429	[2e-01]
	β	-1.774	-0.024	[4e-03]	-0.023	[2e-03]	-0.026	[1e-03]	-0.028	[1e-03]
	ρ	0.600	-0.016	[3e-02]	-0.032	[2e-02]	-0.038	[9e-03]	-0.037	[5e-03]
	ATEF	1.571	-4e-02	[3e-03]	-4e-02	[2e-03]	-4e-02	[2e-03]	-4e-02	[2e-03]
	ATEFX	-2.788	4e-02	[1e-02]	4e-02	[8e-03]	4e-02	[4e-03]	3e-02	[3e-03]
(II)	λ	0.500	-6e-02	[4e-03]	-6e-02	[4e-03]	-6e-02	[4e-03]	-6e-02	[4e-03]
	Cons	-0.423	-0.003	[2e-02]	-0.011	[1e-02]	-0.003	[5e-03]	0.001	[3e-03]
	β	-1.774	-0.122	[2e-02]	-0.118	[2e-02]	-0.120	[1e-02]	-0.122	[2e-02]
	ρ	0.600	-	-	-	-	-	-	-	-
	ATEF	1.571	-3e-02	[2e-03]	-3e-02	[1e-03]	-3e-02	[1e-03]	-3e-02	[8e-04]
	ATEFX	-2.788	-1e-01	[3e-02]	-1e-01	[2e-02]	-1e-01	[2e-02]	-1e-01	[2e-02]
(III)	λ	0.500	1e-02	[3e-04]	1e-02	[2e-04]	1e-02	[2e-04]	1e-02	[2e-04]
	Cons	-0.423	0.356	[1e-01]	0.347	[1e-01]	0.350	[1e-01]	0.354	[1e-01]
	β	-1.774	-0.081	[9e-03]	-0.077	[7e-03]	-0.079	[7e-03]	-0.080	[7e-03]
	ρ	0.600	-	-	-	-	-	-	-	-
	ATEF	1.571	3e-02	[2e-03]	3e-02	[1e-03]	2e-02	[9e-04]	2e-02	[8e-04]
	ATEFX	-2.788	-2e-01	[4e-02]	-2e-01	[3e-02]	-2e-01	[3e-02]	-2e-01	[3e-02]

[a] Selection rate 0.8. (I) Bias correction with dynamic network. (II) No bias correction with dynamic network. (III) No bias correction with fixed network. 500 Monte Carlo simulations.

2.4.2 Exclusion restriction

In this subsection, we demonstrate the importance of having the exclusion restriction condition satisfied. Following the setup in the previous section, we conduct 500 Monte Carlo under DGP(i) with $\lambda = 0.05$ and average selection probability is 0.8. We compare the Monte Carlo standard deviation of the estimator with the case when there is no exclusion restriction, i.e. there is only one covariate heatlh in the selection stage. We report results for the second stage only.

Table 2.6: Comparison of variance of the two-stage estimator without and with exclusion restriction

		ng = 20		ng = 40		ng = 80		ng = 160	
		Bias	Sd	Bias	Sd	Bias	Sd	Bias	Sd
Without exclusion restriction									
λ	0.050	1e-03	(0.018)	-7e-04	(0.013)	3e-05	(0.009)	-6e-04	(0.006)
Cons	-0.423	0.026	(0.512)	-0.028	(0.370)	-0.008	(0.256)	0.002	(0.183)
β	-1.774	-0.007	(0.132)	0.007	(0.097)	0.003	(0.067)	-0.001	(0.048)
ρ	0.600	-0.027	(0.599)	0.029	(0.443)	0.008	(0.307)	-0.003	(0.219)
ATEF	1.040	1e-03	(0.015)	-4e-04	(0.011)	8e-05	(0.007)	-5e-04	(0.005)
ATEFX	-1.845	-1e-02	(0.151)	8e-03	(0.107)	3e-03	(0.074)	-8e-04	(0.054)
With exclusion restriction									
λ	0.050	-6e-04	(0.018)	-3e-04	(0.012)	-2e-04	(0.009)	-2e-04	(0.006)
Cons	-0.423	-0.009	(0.169)	-0.004	(0.114)	0.001	(0.079)	0.004	(0.060)
β	-1.774	0.001	(0.050)	0.001	(0.034)	0.000	(0.023)	-0.002	(0.018)
ρ	0.600	0.011	(0.142)	0.001	(0.098)	-0.004	(0.070)	-0.003	(0.048)
ATEF	1.040	-3e-04	(0.015)	-2e-04	(0.010)	-1e-04	(0.008)	-2e-04	(0.005)
ATEFX	-1.845	1e-03	(0.060)	2e-03	(0.041)	3e-04	(0.029)	-1e-03	(0.022)

As expected, with exclusion restrion, the overall standard errors are lower than those without exclusion restriction. Exclusion restriction does not seem to affect the variance for endogenous peer effects λ but helps a lot with β and ultimately the average total effects of x.

2.4.3 Performance of bootstrap

We now turn to evaluate the performance of block bootstrap. We conduct 500 Monte Carlo simulations under DGP(i) with $\lambda = 0.05$ and average selection probability is 0.8. We consider the case without exclusion restriction. Bootstrapping is done as following: for each Monte Carlo, we perform 500 bootstraps by drawing at subnetwork level with replacement. We report the finite biases for each coefficient, the MC standard deviations, as well as the average coverage rate of the 95% CI, which is the percentage of Monte Carlo simulations in which the 95% CI constructed by bootstrap contains the true parameter value. The 95% CI is constructed under normality assumption based on the bootstap standard errors. The coverage rate is calculated individually for each parameter.

Overall the bootstrap CI works sufficiently well. The coverage rates are slightly lower, at around 93-94%. One reason could be due to the fact that blocks vary significantly in terms of size. In our pool of subnetworks, the minimum size of subnetwork is 1, the maximum is 513, the mean is 184 while the standard deviation of subnetwork size is 116.5.

Table 2.7: Finite sample biases and performance of block bootstrap

		ng = 20			ng = 40			ng = 80			ng = 160		
		Bias	Sd	Coverage	Bias	Sd	Coverage	Bias	Sd	Coverage	Bias	Sd	Coverage
Selection stage: NFP MLE													
δ	0.200	-6e-04	(0.075)	[0.932]	-7e-03	(0.052)	[0.940]	-1e-03	(0.037)	[0.944]	4e-04	(0.026)	[0.942]
Cons	-0.522	0.001	(0.076)	[0.946]	0.002	(0.056)	[0.936]	0.000	(0.038)	[0.956]	-0.001	(0.028)	[0.930]
γ	0.646	-0.001	(0.035)	[0.916]	0.002	(0.023)	[0.934]	0.001	(0.016)	[0.954]	0.001	(0.012)	[0.936]
Outcome stage: NLS													
λ	0.050	-1e-03	(0.019)	[0.920]	4e-04	(0.013)	[0.954]	-2e-05	(0.009)	[0.954]	9e-06	(0.007)	[0.922]
Cons	-0.423	-0.042	(0.516)	[0.938]	-0.006	(0.368)	[0.940]	-0.003	(0.255)	[0.950]	0.008	(0.187)	[0.946]
β	-1.774	0.009	(0.134)	[0.926]	0.003	(0.096)	[0.926]	0.001	(0.067)	[0.940]	-0.002	(0.049)	[0.942]
ρ	0.600	0.049	(0.619)	[0.936]	0.007	(0.432)	[0.948]	0.003	(0.302)	[0.950]	-0.006	(0.224)	[0.944]

500 Monte Carlo simulations with 500 bootstraps for each MC.

2.5 Conclusion

The allowing for network dynamics by sample selection can be useful in different settings in economics. For instance, it can be seen as a type of heterogeneous peer effects where the dimension of heterogeneity is a direct result of choices made by agents in previous periods. Typically researchers divide the peers into separate groups according to some exogenous variables, such as gender or race so that the network can still be considered exogenous (Patacchini et al., 2017; Hsieh and van Kipperluis, 2018). In the particular model specification in this chapter, heterogeneity takes the simplest form possible due to sample selection: shutting down all channels of peer effects of non-selecting individuals. However, the exact form of the network matrix in the second stage can be other non-linear transformation of the original network and choices in the first stage. Yet more guidance from micro-theory is needed to motivate such transformation.

In the model of peer effects with sample selection in this dissertation, we do not deal with network evolution directly. In fact, we do not observe the network overtime but only assume its change through the selection process. Nevertheless, we show that even if a person is endowed with exogenous friendships, the (similar) choices and actions that an individual takes due to this original network would likely change the nature and status of connections. The model here therefore captures a type of indirect relationship between outcome and networks via the selection decision. It is of particular interest to allow for the direct feedback between network links and outcomes: once people linked together, they tend to acquire similar characteristics and values, which in turn has consequences for the whole network itself. This line of modeling belongs to the literature of adaptive networks which describes the coevolution of network and behaviors of members/nodes in the network. Adaptive networks have been studied in physics, chemistry, biology, sociology, and in network sciences in general (Snijders et al., 2007). In Economics, there have been some works that model and estimate the feedback loop of network and outcomes such as those of Boucher (2016); Badev (2017) who study continuous outcome and binary outcome, respectively. Censored outcomes in the type of Tobit-II as considered in our paper would benefit greatly from a similar analysis.

Chapter 3: Peer Effects among High School Students' Subject GPAs in Add Health Data

3.1 Introduction

In this chapter, we apply the methodology proposed in Chapter 2 to an actual data set. In particular, we use The National Longitudinal Study of Adolescent and Adult Health (Add Health), which is a longitudinal study of a nationally representative sample of adolescents in grades 7-12 in the United States in 1994-1995. We wish to measure peer effects among students' subject GPAs, in Social Studies and Science, allowing for potential endogenous selection to take course in the subject. The GPAs in Science and Social Studies are considered as two separate choices and modeled independently.

In our data, we find limited evidence for peer effects in the subject selection decision. While the endogenous selection is not statistically significant, the peer effects coefficients in the subject-specific GPA outcomes are significant. We present the marginal effects which capture the spillovers in the network.

3.2 Add Health data and background context of high school graduation requirements

In their course selection, students in the United States are subject to different high school graduation requirements concerning the number of credits in each subject. In 1983, the National Comission on Excellence in Education produced an influential report, titled *A Nation at Risk*, that expressed concern with the educational quality in the United States. The report criticized the amount of freedom allowed in selecting high school coursework which resulted in a less demanding curriculum, such that 25% of the credits were in electives (PE, work experience, personal and development courses) and remedial coursework. One recommendation of the report is that that states raise high school graduation requirements in core subjects of English, Math, Social Studies and Science. In response to this report, many states have legislated increases in their minimum high school graduation requirements over the last several decades. In the late 80s and 90s, the average credits required across states are 3.5 for English, 2.5 for Social Studies, 2 for Math, and 1.8 for Science. These numbers increase to 3.7, 2.8, 3.0, and 2.7 respectively in 2013.[1]

There is a rich literature in education looking into the effect of state high school graduation requirements, especially in Math and Science, on the subsequent outcomes of students such as college attendance, major, future earnings, as well as health outcomes (Levine and Zimmerman, 1995; Schiller and Muller, 2003; Teitelbaum, 2003; Federman, 2007; Goodman, 2019; Hao and Cowan, 2019; Mangrum, 2019). The overall findings suggest that state minimum requirements increase the number of basic level courses taken by students, especially lower-skilled students from more disadvantaged backgrounds. Whether such courses subsequently help to expose students to higher-level courses that are important in college attendance, in major and occupation choice, is still debatable. At the same time, the general perception is that these state policies would have little impacts on high-skilled students, as these students already exceed such graduation requirements on their own by choice.

When we take into account peer effects, however, the impacts of a policy targeted on a subgroup of students may have spillovers to other students in the same school. If the type of connections that matters for performance in Science in high school networks are unchanged,[2] then there is no change in the exposure among higher and lower-achieving students and therefore there might be no spillovers. However, if we assume dynamic networks, then the hypothetical requirement of all students taking courses in Science may increase exposure to negative shocks to all members of the school network.[3] In the next section, we describe a model that belongs to the latter scenario. In our setting, the original friendship structure within a school is assumed to be exogenous. The decision to take a course in a subject by students can effectively influence the type of peers who matters for the subject performance: only students taking courses in the same subject can influence each other. If a school policy requires all students to take courses in the subject, students will now be influenced also by their original friends who would not have chosen the

[1]Caculation based on Digest of Education Statistics data.

[2]If networks are described by strong homophily in terms of academic achievement (higher-achieving students tend to befriend each other), or it may be the case that schools place students of similar ability in the same classroom, which effectively fix the network structure.

[3]On the other hand, if education is effective and therefore students can accumulate and improve cognitive capital, then such policy may help to reduce the distance among individuals. In a model with fully dynamic networks where members actively change their connection statuses, new links may be formed.

subject otherwise.

3.3 Data

In this section we describe more in details the samples we obtain from the full Add Health survey.

There are two important features of Add Health study that are especially attractive to researchers in networks. First, in the first wave, conducted from September 1994 - April 1995, Add Health administered a special survey in which all students in 144 participating schools filled out a questionnaire. Second, the questionnaire elicited detailed network data, asking students to nominate up to ten friends, five in each gender. While some students did nominate friends outside of schools, more than 75% of the connections occur within schools. As a result, it is possible to match these nominees from the pool of completed in-school questionnaires. The design therefore gives rich and saturated information about social networks within each school. What makes Add Health relevant to our interest is that the survey further collects basic demographic information as well as academic performance and health outcomes. Particularly, Add Health asks for GPA in four different subjects English/Language Arts, Mathematics, History/Social Studies, and Science. The higher rate of missing response to Social Studies and Science suggests there might be a place for correcting for endogenous sample selection.

We focus on the estimation of peer effects within each of the two subjects: Science and Social Studies. There are two dependent variables. The first variable is the selection indicator specifying whether students took a Science course in the most recent grading period. Students are considered to have taken Science courses if they report their letter grades. They are considered to not have taken any Science courses if they report explicitly so in the questionnaire. The second variable is a discrete outcome $1, 2, 3, 4$ corresponding to D (or lower), C, B, A if students did take Science courses, and missing otherwise. We drop the observations if the selection indicator is missing, i.e. if missing GPA is due to other types.

We include the following set of control variables on students' basic demographic information: age, gender, grade, Hispanic, race (white, black, Asian, other races), number of years in school, health (1-5, 5 is lowest), indicators for living with mother, with father; mother and father education (less than high school, high school, college), mother and father occupation (homemaker, professionals, office workers and sales, manual and farming, military, and "do not know"), and household size.[4] Since we assume the original friendship within a school to be exogenous, we include many variables to account for the friendship formation process. Among many factors that influence the probability of linking, students' ability could play a central role and therefore is proxied by the inclusion of parents' education and occupation dummies. Sample selection, in this case, should be viewed as the unobserved random shock to students' potential performance in the subject that is unrelated to ability.

Due to the extensive nature of the Add Health survey, actual information on states where schools locate is not available and therefore obtaining the specific high school graduation requirements for each school is not possible for confidentiality. To account for school characteristics, we include indicators for size, for school type, region (West, Midwest, South, NorthEast), and urbanicity (Rural, Suburban, Urban).

In terms of network structure, we consider each school to be an independent network and keep only the links in which both the nominator and the nominee are in the same school. We restrict the study to high school students in grades 9 and above. We stress again that missing data is a substantial challenge for network models, more so than traditional data consisting of independent observations. The reason is that removing observations due to any missing variables means the removal of members and connections in the network. To make sure that we have "acceptable" network coverage, we compare the number of observations we have in each school (after removing individuals with missing variables) with the total school size obtained from administrative data. We keep only schools where there is at least 50% of students that can be found in our sample.

The original number of students completing the questionnaire is more than 85,000, among which the number of students from grade 9 and above is around 58,000. Our final

[4]Students are asked whether they live with mother or father. If the answer is yes then information about the corresponding parent's education and occupation is solicited. Therefore, we drop observations with missing value in education or occupation only if missing is not due to this type of valid skip.

sample size is about 27,000 from 44 schools. In Appendix **B.1**, we provide the summary statistics for variables at the individual level and at the school level.

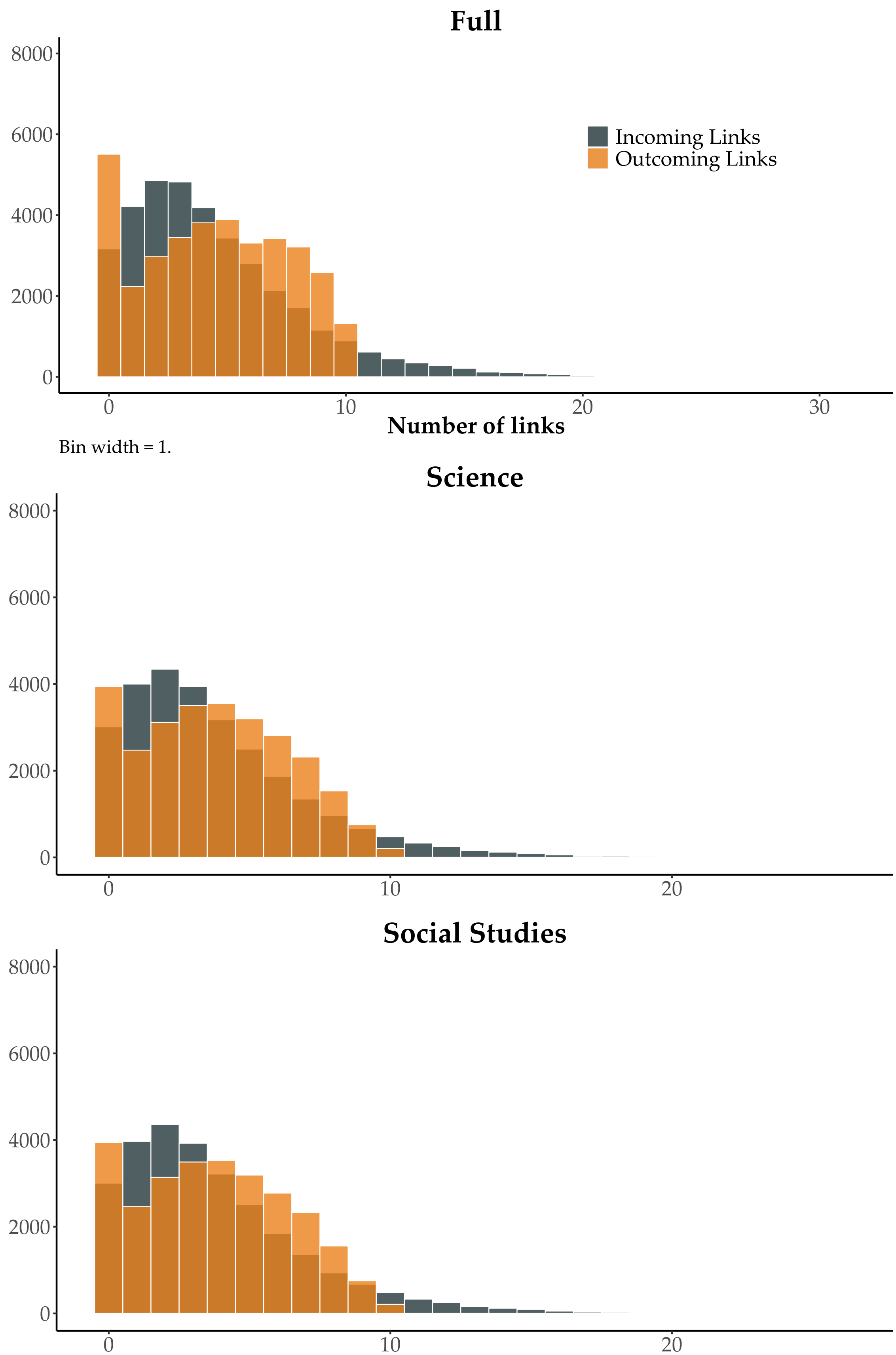
A potential concern is the network structure in our samples. To see how our final network structures compared to before individuals are removed due to missing variables, we collect the original members and nominations in the schools that show up in our samples. In Table **3.1**, we report basic network statistics when we consider all schools as a large network. Viewed in this way, the large network is very sparse, with density ranging from 0.012% to 0.014%. The network degree (average connections per member) does decrease in our samples compared to the original network structure. This seems to come principally from the removal of some member with high in-degree (number of incoming links i.e. number of times being nominated). Fig. **3.1** illustrates the distributions of links in all samples. Other features of the large network do not differ significantly in different samples. The proportion of isolated students and the proportion of reciprocal connections (both students nominate each other) in all samples are quite similar, at around 4% and 40% respectively across samples. We also look into the characteristics of each subnetwork (school) and report the average across these subnetworks. From this angle, network density averages around 0.8-0.9%. These subnetworks are still sparse.

Table 3.1: Network statistics

	Full	Science	Social Studies
Nodes - Roster	46348	46348	46348
Nodes - Sampled	35722	27408	27392
Nodes - Isolated	1204	1071	1086
(Proportion)	0.034	0.039	0.040
Coverage	0.77	0.59	0.59
Links	159181	103612	103580
Links - Reciprocal	64614	43990	43956
(Proportion)	0.41	0.42	0.42
Density (x100)	0.012	0.014	0.014
Degree	4.46	3.78	3.78
In-degree (sd)	3.61	3.19	3.18
In-degree (min)	0	0	0
In-degree (max)	31	26	26
Out-degree (sd)	3.03	2.61	2.61
Out-degree (min)	0	0	0
Out-degree (max)	10	10	10

Full: The original network structure before dropping observation due to missing variables

Figure 3.1: Links distribution



3.4 Estimation

In this section we summarize the model and the estimation procedure in Chapter 2. The model comprises of Eq. (2.2.2) that describes the decision to take course in a subject

$$p_i^* = Pr\left(\underbrace{\mathbf{z}_i\gamma + \delta\mathbf{g}_i^*\mathbf{p}_n^*}_{r_i} + v_i > 0\right) \quad i = 1, \ldots, n$$
$$d_i = 1\left(v_i > -\mathbf{z}_i\gamma - \delta\mathbf{g}_{i,n}^*\mathbf{p}_n^*\right)$$

and Eq. (2.2.4) that describes the realization of subject-specific GPAs among students taking the same subject only

$$y_i^* = \lambda\mathbf{g}_i^{o*}\mathbf{y}_n^* + \mathbf{x}_i\boldsymbol{\beta} + u_i$$
$$\mathbf{y}_n^* = \lambda G_n^{o*}\mathbf{y}_n^* + \mathbf{X}_n\boldsymbol{\beta} + \mathbf{u}_n$$

We have δ captures the endogenous peer effects in the first stage and λ captures the endogenous peer effects in the second stage. G_n^{o*} is the network structure (row-normalized) realized in the subject GPAs realization such that all links to and from individuals not choosing to take course in the subjects are set to zero.

In the first step estimation, we use the nested fixed point estimator to solve for equilibrium inside the maximization of the log likelihood:

$$\max_{\boldsymbol{\theta}^d\in\boldsymbol{\Theta}^d} \frac{1}{n}\sum_i \left(d_i p_i^* + (1-d_i)(1-p_i^*)\right) \tag{3.4.1}$$

where under parametric assumption that the error term v follows standard normal distribution. In other words, we solve for the equilibrium $\mathbf{p}_n^*$ by iteration for each search of parameter values.

To estimate the second stage, we construct a bias correction term from another parametric assumption about the relationship between u and v. We assume that u and v are distributed as bivariate normal with correlation $\tilde{\rho}$ and variance of error in second stage

σ_u^2. This means that $E(u_i \mid d_i) = \underbrace{\tilde{\rho}\sigma_u}_{\rho} E(v_i \mid d_i) = \rho b_i(\boldsymbol{\theta}^d)$ of which closed form is easily calculated from a truncated normal distribution, as with Heckman sample selection term. Because:

$$E[\mathbf{u}_n \mid \mathbf{d}_n] = \begin{pmatrix} E(u_1 \mid \mathbf{d}_n) \\ \vdots \\ E(u_n \mid \mathbf{d}_n) \end{pmatrix} = \rho \begin{pmatrix} E(u_1 \mid d_1) \\ \vdots \\ E(u_n \mid d_n) \end{pmatrix} = \rho \begin{bmatrix} b_1(\boldsymbol{\theta}^d) \\ \vdots \\ b_n(\boldsymbol{\theta}^d) \end{bmatrix}$$

so that we can write $\hat{E}(\mathbf{u}_n \mid \mathbf{d}_n, G_n) = \rho \mathbf{b}_n(\hat{\boldsymbol{\theta}}^d)$ and plug in to estimate the following equation by non-linear least squares to estimate Eq. (2.3.3):

$$\min_{\boldsymbol{\theta}^y \in \boldsymbol{\Theta}^y} \frac{1}{n} \left(\mathbf{y}_n^1 - J_n^1 (I_n - \lambda G_n^{o*})^{-1} \left[\mathbf{X}_n \boldsymbol{\beta} + \rho \mathbf{b}_n(\widehat{\boldsymbol{\theta}}^d)\right]\right)' \left(\mathbf{y}_n^1 - J_n^1 (I_n - \lambda G_n^{o*})^{-1} \left[\mathbf{X}_n \boldsymbol{\beta} + \rho \mathbf{b}_n(\widehat{\boldsymbol{\theta}}^d)\right]\right) \tag{3.4.2}$$

Inference is done via block bootstrap by drawing schools with replacement in 500 bootstraps.

3.5 Results

Our results for main coefficients of interest, 'endogenous' peer effects for selection stage δ, 'endogenous' peer effects for outcome stage λ, and strength of selection ρ, are summarized in Table **3.2**. The full results can be found in Appendix **B.2**.

In Table **3.2**, we provide estimates from three different specifications. (I) is our main 2-step estimator that corrects for potential bias from sample selection and assumes dynamic peer structures (i.e. the endogenous peer effects come from selected friends only). (II) and (III) are NLS estimators for the second stage of GPA outcomes that assume random selection (does not correct for sample selection bias). (II) assumes dynamic peer structure but (III) assumes non-dynamic peer structure. Standard errors for (I), (II), (III) are obtained by 500 bootstraps. Table **3.2** does not consider contextual peer effects.

In general, we do not have evidence for selection. In other words, students' choice of taking a course in a subject does not seem to correlate with their subject-specific outcomes. In the appendix, we provide estimates by the traditional Heckit which agree with this result.

Table 3.2: Estimation results for main parameters

		Main (I)		(II)		(III)	
Science	Peer effects selection (δ)	0.144*	(0.056, 0.231)	-	-	-	-
	Peer effects outcome (λ)	0.053*	(0.035, 0.070)	0.050*	(0.037, 0.063)	0.050*	(0.037, 0.064)
	Selection (ρ)	0.243	(-0.658, 1.145)	-	-	-	-
Social Studies	Peer effects selection (δ)	0.070	(-0.009, 0.149)	-	-	-	-
	Peer effects outcome (λ)	0.061*	(0.041, 0.081)	0.053*	(0.036, 0.071)	0.062*	(0.044, 0.079)
	Selection (ρ)	1.130	(-0.401, 2.662)	-	-	-	-

(I): Bias correction with dynamic network. (II): No bias correction with dynamic network. (III): No bias correction with fixed network. 95% CI in parentheses, computed from 500 bootstraps. * denotes CI does not include 0.

The 'endogenous' peer effect in the selection into Science class is statistically significant. This means that there is some dependency in friends' selection probabilities. Heuristically, we can consider a typical student whose covariates take the mean value in the sample. Holding everything else constant, if the average of friends' probabilities of selection into Science courses shifts from 0 to 50%, then the probability that this student also takes a course in Science would increase by 1.36%. The marginal effects of friends' selection probabilities are practically very small. This is understandable, given that the average selection in our samples are high, around 88-90%. The 'endogenous' peer effects in the subject-specific outcome is around 0.05 in both subjects. Similarly heuristically, this means, holding everything else constant, a 1 point increase in friends' average GPA would increase own's GPA by around 0.05-0.06 point. 'Holding everything else constant' means we do not yet take into account the feedback from own's GPA to the GPAs of other members in the network.

In Tables 3.3 and 3.4, we add contextual peer effects. The full results are reported in Appendix B.2. Since we have two different types of peer effects now, it is important to discuss which type of peer effects follows which peer structure. We now present six possible configurations. The first set of estimators, (I).a, (I).b, and (I).c, corrects for sample selection bias. The second set of estimators, (II).a, (II).b, and (III).c, do not. The subject selection decision is modeled the same for all three estimators (I), (II), (III), where the endogenous peer effects and the contextual peer effects come from the same peer structure - the original friendship structure. In other words, they all assume the following first stage:

$$p_i^* = Pr\left(\mathbf{z}_i\gamma^{own} + \delta\mathbf{g}_i^*\mathbf{p}_n^* + \mathbf{g}_i^*\mathbf{Z}_n\gamma^{contextual} + v_i > 0\right) \quad i = 1, \ldots, n$$

In the second stage of subject GPA, the two peer structures need not be identical. In (I).a (and (II).a), the endogenous peer effects occur among the selected friends only, while the contextual peer effects come from all original friends:

$$\mathbf{y}_n^* = \lambda G_n^{o*}\mathbf{y}_n^* + G_n^*\mathbf{X}_n\boldsymbol{\beta}^{contextal} + \mathbf{X}_n\boldsymbol{\beta}^{own} + \mathbf{u}_n$$

In (I).b (and (II).b), both the endogenous peer effects and the contextual peer effects come from the selected friends.

$$\mathbf{y}_n^* = \lambda G_n^{o*}\mathbf{y}_n^* + G_n^{o*}\mathbf{X}_n\boldsymbol{\beta}^{contextal} + \mathbf{X}_n\boldsymbol{\beta}^{own} + \mathbf{u}_n$$

In (I).c (and (III).c), both the endogenous peer effects and the contextual peer effects come from the original friends:

$$\mathbf{y}_n^* = \lambda G_n^*\mathbf{y}_n^* + G_n^*\mathbf{X}_n\boldsymbol{\beta}^{contextal} + \mathbf{X}_n\boldsymbol{\beta}^{own} + \mathbf{u}_n.$$

The bias correction in this paper is designed for (I).a and (I).b, but it can still be useful for (I).c.[5]

[5] See the discussion on page 13 for the required bias correction for (I).c. Even though our estimator is not designed for (I).c in mind, the bias correction term we propose can still be useful since all other additive terms for individual i in $S_n^{-1}(\lambda)_{\text{ith row}}E(\mathbf{u}_n \mid \mathbf{d}_n)$ can be viewed as random noises. Our simulations support this argument.

Table 3.3: Estimation results for main parameters with contextual peer effects - Bias correction

		(I).a		(I).b		(I).c	
Science	Peer effects selection (δ)	0.592	(-0.726, 1.909)	0.592	(-0.726, 1.909)	0.592	(-0.726, 1.909)
	Peer effects outcome (λ)	0.102*	(0.050, 0.155)	0.463*	(0.340, 0.587)	0.486*	(0.329, 0.644)
	Selection (ρ)	-0.337	(-1.226, 0.553)	0.100	(-0.532, 0.733)	-0.001	(-0.303, 0.301)
Social Studies	Peer effects selection (δ))	2.223*	(0.440, 4.006)	2.223*	(0.440, 4.006)	2.223*	(0.440, 4.006)
	Peer effects outcome (λ)	0.036	(-0.020, 0.092)	0.392*	(0.258, 0.527)	0.436*	(0.296, 0.576)
	Selection (ρ)	-0.307	(-1.072, 0.458)	-0.073	(-0.610, 0.464)	-0.094	(-0.314, 0.126)

(I).a: dynamic peer structure for endogenous peer effect, fixed peer structure for contextual peer effects. (I).b: dynamic peer structure for both types of peer effects. (I).c: fixed peer structure for both types of peer effects. 95% CI in parentheses, computed from 500 bootstraps for Science and 250 bootstraps for Social Studies. * denotes CI does not include 0. The endogenous peer effect in selection δ is constrained within $(-\sqrt{2\pi}, \sqrt{2\pi})$ in the maximum likelihood function.

Table 3.4: Estimation results for main parameters with contextual peer effects - No bias correction

		(II).a		(II).b		(III).c	
Science	Peer effects outcome (λ)	0.107*	(0.057, 0.156)	0.461*	(0.338, 0.583)	0.486*	(0.354, 0.618)
Social Studies	Peer effects outcome (λ)	0.045	(-0.006, 0.095)	0.399*	(0.272, 0.526)	0.434*	(0.318, 0.550)

(II).a: dynamic peer structure for endogenous peer effect, fixed peer structure for contextual peer effects. (II).b: dynamic peer structure for both types of peer effects. (III).c: fixed peer structures for both types of peer effects. 95% CI in parentheses, computed from 500 bootstraps for Science and 250 bootstraps for Social Studies. * denotes CI does not include 0.

As we include these contextual effects, the optimization has a much more difficult time in obtaining the unique equilibrium selection probabilities in the estimation of the first stage. We conduct 500 bootstraps for Science sample and 250 bootstraps for Social Studies sample. The endogenous peer effects in the selection in Science has a very noisy estimate and is not statistically significant. A few contextual peer effects are significant. In the sample for Social Studies, the estimate almost reaches the upper boundary of its parameter space.[6] 70% of the 250 bootstraps for Social Studies give estimates for $\delta > 2.4$. As a result, we focus on the results for Science.

The estimate for selection term is still noisy and statistically zero. It seems that sample selection is not likely a feature in our data. However, there is an interesting observation for the endogenous peer effects in the second stage. With contextual peer effects, the strength of the endogenous peer effects (λ) in the subject GPAs is sensitive to the configuration of which network structures we assume for which types of peer effects. The endogenous peer effects in the subject GPAs are much smaller when we assume two different peer structures for these two types of peer effects. The larger estimates of the endogenous peer effects when assuming the same peer structures for two types of peer effects are of the similar magnitude to the results obtained by Lin (2010) and Hoshino (2017) who study students' overall GPAs with network at school-grade levels. Note that Hoshino (2017) also consider the sample selection issue but proposes a specific bias correction for (I).c. Nevertheless, in our case, sample selection raises a question about which peer structures are relevant for which types of peer effects? The answer ultimately requires a discussion on the fundamental microeconomic theory on the nature of peer effects. Fruehwirth (2014) presents an interesting discussion where peer effects occur among students' unobserved efforts. In sample selection setting, we can argue that efforts for a specific course are only observed among friends taking the same course. The unobserved nature of efforts yields a reduced form of GPA that is a function of both friends' GPAs and friends' contextual characteristics that correspond to the same peer structure. Such theoretical equation could motivate the configuration as in (I).b.[7]

[6] CI is constructed under normality assumption, using standard errors approximated by bootstrap standard errors. This is why the upperbound of the peer effects in selection stage goes to above 4.

[7] We heuristically do a simulation experiment to examine this remarkedly different estimates due to

3.5.1 Marginal effects

As mentioned before, reporting the full $n \times n$ matrix of marginal effects for each exogenous variable is not desirable. Assuming that selection decisions remain fixed, we report two types of marginal effects with respect to the subject GPAs for our main specification (I): (1) Average of direct/own effects (average of diagonal elementstimes β_k); (2) Average of total effects (average of column sums times β_k), which is the average of effects on everyone in the network (including self).

We calculate average effects within each school and average again across all 44 schools. More importantly, we report these marginal effects for the subset of students who actually take course in the subject only. Tables B.12 and B.15 provide the full marginal effects of all exogenous variables along with the OLS estimates.

In the following Table 3.5, we report the multipliers only, i.e. assuming $\beta_k = 1$ for the main estimation results (I) without contextual peer effects:

Table 3.5: Marginal effects multipliers: assuming $\beta_k^{\text{own}} = 1$ for model without contextual peer effects

	Average own effects		Average total effects	
Science	1.000	(1.000, 1.000)	1.047	(1.031, 1.064)
Social Studies	1.000	(1.000, 1.001)	1.054	(1.035, 1.073)

Marginal effects for students selecting the subject only. Marginal effects average within in each subnetwork then average over subnetworks. 95% CI in parentheses, computed from 500 bootstraps.

We can see that the average total effects are larger than the own effects. The spillovers, however, are small, at around 4-5% of the own effects. The endogenous peer effects λ therefore could be interpreted as the overall spillover level in the network.

When we include contextual peer effects, the marginal effects are calculated instead from two matrices $(I_n - \lambda G_n^{*\text{endo}})^{-1}\beta_k^{\text{own}}$ and $(I_n - \lambda G_n^{*\text{endo}})^{-1} G_n^{*\text{context}} \beta_k^{\text{context}}$ where $G_n^{*\text{endo}}$ is the peer structure for the endogenous peer effects and $G_n^{*\text{context}}$ is the peer structure for the contextual peer effects. We report these marginal effects separately for own characteristics

different network structure assumption in Appendix B.3. Simulation results seem to support (I).c or (III).c.

and for contextual characteristics. For simplicity, assume $\beta_k^{\text{own}} = \beta_k^{\text{context}} = 1$, so that we report the multipliers only. The summary is in Table 3.6.

Table 3.6: Marginal effects multipliers for models (I).a, b, c with contextual peer effects Assuming $\beta_k^{\text{own}} = \beta_k^{\text{context}} = 1$

		$(I_n - \lambda G_n^{*\text{endo}})^{-1}$				$(I_n - \lambda G_n^{*\text{endo}})^{-1} G_n^{*\text{context}}$			
		Average own effects		Average total effects		Average own effects		Average total effects	
Science	(I).a	1.001	(1.000, 1.002)	1.096	(1.041, 1.152)	0.010	(0.005, 0.015)	0.871	(0.823, 0.918)
	(I).b	1.027	(1.010, 1.045)	1.691	(1.349, 2.032)	0.059	(0.038, 0.080)	1.490	(1.172, 1.808)
	(I).c	1.029	(1.007, 1.050)	1.698	(1.279, 2.117)	0.059	(0.034, 0.084)	1.436	(1.045, 1.826)
Social Studies	(I).a	1.000	(0.999, 1.001)	1.031	(0.979, 1.083)	0.003	(-0.002, 0.009)	0.787	(0.744, 0.829)
	(I).b	1.019	(1.003, 1.035)	1.516	(1.209, 1.822)	0.049	(0.027, 0.070)	1.315	(1.032, 1.598)
	(I).c	1.022	(1.006, 1.038)	1.555	(1.248, 1.861)	0.051	(0.030, 0.071)	1.272	(0.988, 1.557)

Marginal effects for students selecting the subject only. Marginal effects average within in each subnetwork then average over subnetworks. 95% CI in parentheses, computed from 500 bootstraps for Science and 250 bootstraps for Social Studies.

The relative size of average total effect compared to average own effect is larger as λ increases, as expected. Marginal effects of each covariate when using the appropriate point estimates are quite different for each configuration. The signs for statistically significant covariates are in general kept throughout, however.

3.6 Conclusion

In this chapter, we study sample selection with peer effects where network structure in the outcome stage is influenced by choices in the selection stage. We apply the estimation method to study high school students' subject-specific GPAs in Social Studies and Science. We find some 'endogenous' peer effects in the decision to take courses in Science and Social Studies as well as in the academic performance recorded by GPA for each subject. However, we do not detect a statistically significant endogenous selection mechanism at the level of the aggregate outcome obtained. Our results suggest that, with sample selection situation, researchers interested in both endogenous and contextual peer effects have a close look on the possibility of the difference in peer structures for different types of peer effects, at least in the linear spatial models studied here.

Chapter A: Appendix for Chapter 2

A.1 Outcome stage: Incomplete information

We can argue that GPA is neither fully controlled by students nor fully observed by friends. It is possible that students can observe friends' ability for which GPA is a proxy. In that case we can define an incomplete information setting such as the following. As students start taking classes, in addition to knowledge on G_n and $\mathbf{Z}_n$, now they also know about their friends' selection $\mathbf{d}_n$ and the mathematical expectation is taken conditional on this information set. To simplify notation, we ommit $G_n, \mathbf{Z}_n$ from the conditioning set.

$$\begin{aligned} y_i^* &= \lambda \omega_i \sum_j g_{ij} d_j E(y_j^* \mid \mathbf{d}_n) + \mathbf{X}_i \boldsymbol{\beta} + u_i \\ E(y_i^* \mid \mathbf{d}_n) &= \lambda \omega_i \sum_j g_{ij} d_j E(y_j^* \mid \mathbf{d}_n) + \mathbf{X}_i \boldsymbol{\beta} + E(u_i \mid \mathbf{d}_n). \end{aligned} \tag{A.1.1}$$

We can solve for the reduced form of $E(y_i^* \mid \mathbf{d}_n, \mathbf{Z}_n, G_n)$:

$$\begin{aligned} E(\mathbf{y}_n^* \mid \mathbf{d}_n) &= \lambda G_n^{o*} E(\mathbf{y}_n^* \mid \mathbf{d}_n,) + \mathbf{X}_n \boldsymbol{\beta} + E(\mathbf{u}_n \mid \mathbf{d}_n) \\ &= (I_n - \lambda G_n^{o*})^{-1} \left[\mathbf{X}_n \boldsymbol{\beta} + E(\mathbf{u}_n \mid \mathbf{d}_n) \right] \end{aligned} \tag{A.1.2}$$

Plug (A.1.2) into (A.1.1), we obtain the equilibrium for outcome stage:

$$\begin{aligned} \mathbf{y}_n^* &= \lambda G_n^{o*} E(\mathbf{y}_n^* \mid \mathbf{d}_n) + \mathbf{X}_n \boldsymbol{\beta} + \mathbf{u}_n \\ &= \lambda G_n^{o*} (I_n - \lambda G_n^{o*})^{-1} \left[\mathbf{X}_n \boldsymbol{\beta} + E(\mathbf{u}_n \mid \mathbf{d}_n) \right] + \mathbf{X}_n \boldsymbol{\beta} + \mathbf{u}_n \\ &= (I_n + \lambda G_n^{o*} (I_n - \lambda G_n^{o*})^{-1}) \mathbf{X}_n \boldsymbol{\beta} + \lambda G_n^{o*} (I_n - \lambda G_n^{o*})^{-1} E(\mathbf{u}_n \mid \mathbf{d}_n) + \mathbf{u}_n \\ &= \underbrace{(I_n - \lambda G_n^{o*})^{-1} \mathbf{X}_n \boldsymbol{\beta}}_{B_1} + \underbrace{\lambda G_n^{o*} (I_n - \lambda G_n^{o*})^{-1} E(\mathbf{u}_n \mid \mathbf{d}_n)}_{B_2} + \mathbf{u}_n \end{aligned} \tag{A.1.3}$$

where the last equality in (A.1.3) is due to $(I - A)^{-1} = \sum_{k=0}^{\infty} A^k$ for A be a non-singular square matrix. To correct for bias from sample selection, we modify (A.1.3):

$$\mathbf{y}_n^* = B_1 + B_2 + E(\mathbf{u}_n \mid \mathbf{d}_n) + \underbrace{\mathbf{u}_n - E(\mathbf{u}_n \mid \mathbf{d}_n)}_{\boldsymbol{\epsilon}_n}$$

but then

$$\begin{aligned} B_2 + E(\mathbf{u}_n \mid \mathbf{d}_n) &= (I_n + \lambda G_n^{o*}(I_n - \lambda G_n^{o*})^{-1})E(\mathbf{u}_n \mid \mathbf{d}_n) \\ &= (I_n - \lambda G_n^{o*})^{-1} E(\mathbf{u}_n \mid \mathbf{d}_n) \end{aligned}$$

using the same logic of the last equality in (A.1.3). So that we now have:

$$\begin{aligned} \mathbf{y}_n^* &= (I_n - \lambda G_n^{o*})^{-1} \left[\mathbf{X}_n \boldsymbol{\beta} + E(\mathbf{u}_n \mid \mathbf{d}_n)\right] + \boldsymbol{\epsilon}_n \\ &= S_n^{-1}(\lambda) \left[\mathbf{X}_n \boldsymbol{\beta} + E(\mathbf{u}_n \mid \mathbf{d}_n)\right] + \boldsymbol{\epsilon}_n \end{aligned} \tag{A.1.4}$$

$$J_n^1 \mathbf{y}_n^* = J_n^1 S_n^{-1}(\lambda) \left[\mathbf{X}_n \boldsymbol{\beta} + E(\mathbf{u}_n \mid \mathbf{d}_n)\right] + J_n^1 \boldsymbol{\epsilon}_n \tag{A.1.5}$$

A.2 Mathematical notes

Lemma A.2.1 *If A_n and ${}_nB$ are $n \times n$ uniformly bounded matrices in both row and column sums by c_a, c_b, then so is $D = A_n B_n$.*

we have $d_{ij} = \sum_{r=1}^{n} a_{ir} b_{rj}$. The ith row sum of D is

$$\sum_{j=1}^{n} d_{ij} = \sum_{j=1}^{n} |d_{ij}| \leq \sum_{j=1}^{n}\sum_{r=1}^{n} |a_{ir}||b_{rj}| = \sum_{r=1}^{n} |a_{ir}| \sum_{j=1}^{n} |b_{rj}| \leq c_b \sum_{r=1}^{n} |a_{ir}| \leq c_b c_a$$

Similar for the ith column sum of D.

Lemma A.2.2 *If A_n is $n \times n$ uniformly bounded matrices in both row and column sums and B_n is $k \times n$ matrix of real numbers that are bounded in absolute value ($|b_{ij}| \leq c_b \forall i, j, n$). Then elements of $D = BA$ is also bounded in absolute value.*

$$|d_{ij}| \leq \sum_{r=1}^{n} |b_{ir}||a_{rj}| \leq c_b \sum_{r} |a_{rj}| \leq c_b c_a$$

Lemma A.2.3 *If A_n is $n \times n$ uniformly bounded matrices in both row and column sums and $\mathbf{x}, \mathbf{y}$ are $n \times 1$ column vectors of real numbers that are bounded in absolute value ($|d_i| \leq c_d \forall i, n$). Then $|\mathbf{x}' A_n \mathbf{y}| = O(n)$.*

$$|\mathbf{x}' A_n \mathbf{y}| \leq \sum_j \sum_i |x_i||a_{ij}||y_j| \leq c_d^2 \sum_j \sum_i |a_{ij} \leq n c_d^2 c_a$$

Lemma A.2.4

$$\begin{aligned}
S_n^{-1}(\lambda_0)\mathbf{X}_n\boldsymbol{\beta}_0 &= S_n^{-1}(\lambda)S_n(\lambda)S_n^{-1}(\lambda_0)\mathbf{X}_n\boldsymbol{\beta}_0 \\
&= S_n^{-1}(\lambda)(I_n - \lambda G_n^{o*})S_n^{-1}(\lambda_0)\mathbf{X}_n\boldsymbol{\beta}_0 \\
&= S_n^{-1}(\lambda)S_n^{-1}(\lambda_0)\mathbf{X}_n\boldsymbol{\beta}_0 - S_n^{-1}(\lambda)\lambda G_n^{o*}S_n^{-1}(\lambda_0)\mathbf{X}_n\boldsymbol{\beta}_0 \\
&= S_n^{-1}(\lambda)(I_n + \lambda_0 G_n^{o*}S_n^{-1}(\lambda_0))\mathbf{X}_n\boldsymbol{\beta}_0 - S_n^{-1}(\lambda)\lambda G_n^{o*}S_n^{-1}(\lambda_0)\mathbf{X}_n\boldsymbol{\beta}_0 \\
&= S_n^{-1}(\lambda)\mathbf{X}_n\boldsymbol{\beta}_0 + S_n^{-1}(\lambda)(\lambda_0 - \lambda)G_n^{o*}S_n^{-1}(\lambda_0)\mathbf{X}_n\boldsymbol{\beta}_0 \\
\Rightarrow S_n^{-1}(\lambda_0)\mathbf{X}_n\boldsymbol{\beta}_0 - S_n^{-1}(\lambda)\mathbf{X}_n\boldsymbol{\beta} &= S_n^{-1}(\lambda)\mathbf{X}_n(\boldsymbol{\beta}_0 - \boldsymbol{\beta}) + S_n^{-1}(\lambda)(\lambda_0 - \lambda)G_n^{o*}S_n^{-1}(\lambda_0)\mathbf{X}_n\boldsymbol{\beta}_0 \\
&= S_n^{-1}(\lambda)[\mathbf{X}_n(\boldsymbol{\beta}_0 - \boldsymbol{\beta}) + (\lambda_0 - \lambda)G_n^{o*}S_n^{-1}(\lambda_0)\mathbf{X}_n\boldsymbol{\beta}_0]
\end{aligned}$$

A.2.1 Results for non-stochastic network structure

Lemma A.2.5 *Let A_n be $n \times n$ nonstochastic matrix whose row and column sums are uniformly bounded in absolute value. Let $\mathbf{u}_n' = (u_1, \ldots, u_n)$ where u_i are iid with mean zero and finite variance σ_u^2 and $E(u_i^4) < \infty$, we have the following results*

1. $E(\mathbf{u}_n' A_n \mathbf{u}_n) = \sigma_u^2 tr(A_n) = O(n)$
2. $Var(\mathbf{u}_n' A_n \mathbf{u}_n) = O(n)$
3. $\frac{1}{n}\mathbf{u}_n' A_n \mathbf{u}_n - \frac{1}{n}E(\mathbf{u}_n' A_n \mathbf{u}_n) = o_p(1)$

Mean:

$$\mathbf{u}_n' A_n \mathbf{u}_n = \sum_{j=1}^{n}\sum_{i=1}^{n} u_i a_{ij} u_j = \sum_{i=1}^{n} u_i^2 a_{ii} + \sum_{j<k}\sum u_j u_k (a_{jk} + a_{kj})$$

$E(\mathbf{u}_n' A_n \mathbf{u}_n) = \sum_{i=1}^{n} E(u_i^2) a_{ii} + 0 = \sigma_u^2 \sum_{i=1}^{n} a_{ii} = \sigma_u^2 tr(A_n)$

Variance: $E(u_i^2 u_j u_k) = 0$ for $j \neq k$, so the squared terms are uncorrelated with the double sum terms.

$$\begin{aligned}\mathrm{Var}\left(\mathbf{u}_n' A_n \mathbf{u}_n\right) &= \mathrm{Var}\left(\sum_{i=1}^{n} u_i^2 a_{ii}\right) + \mathrm{Var}\left(\sum_{j<k}\sum u_j u_k (a_{jk} + a_{kj})\right) \\ &= \left(E(u_i^4) - \sigma_u^4\right) \sum_i a_{ii}^2 + E\left[\left(\sum_{j<k}\sum u_j u_k (a_{jk} + a_{kj})\right)^2\right]\end{aligned}$$

Consider:

$$\left(\sum_{j<k}\sum u_j u_k (a_{jk}+a_{kj})\right)^2 = \left(\sum_{j<k}\sum u_j u_k (a_{jk}+a_{kj})\right)\left(\sum_{l<m}\sum u_l u_m (a_{lm}+a_{ml})\right)$$

$$\text{For } j<k, l<m : E(u_j u_k u_l u_m) = \begin{cases} \sigma_u^4 & \text{if } j=l\neq k=m \\ 0 & \text{otherwise} \end{cases}$$

$$E\left[\left(\sum_{j<k}\sum u_j u_k (a_{jk}+a_{kj})\right)^2\right] = \sigma_u^4\left[\sum_{j<k}\sum (a_{jk}+a_{kj})^2\right] = \sigma_u^4\frac{1}{2}\left[\sum_{j\neq k}\sum (a_{jk}+a_{kj})^2\right]$$

$$= \sigma_u^4\frac{1}{2}\left[\sum_j\sum_k (a_{jk}+a_{kj})^2 - \sum_i (2a_{ii})^2\right]$$

Let $M_n = (A_n + A_n')$ is a symmetric matrix with $m_{jk} = m_{kj} = a_{jk} + a_{kj}$.

$$\sum_j\sum_k (a_{jk}+a_{kj})^2 = \sum_j\sum_k m_{jk}^2 = \sum_j\sum_k m_{jk}m_{kj}$$

$$= \sum_j (M_n M_n')_{jj} = tr(M_n^2)$$

$$tr(M_n^2) = tr\left[(A_n + A_n')^2\right] = tr(A_n^2) + tr(A_n' A_n) + tr(A_n A_n') + tr(A_n'^2)$$

$$= 2tr(A_n^2) + 2tr(A_n A_n')$$

Because $tr(A_n A_n) = tr((A_n A_n)') = tr(A_n' A_n')$. In the end:

$$\mathrm{Var}\left(\mathbf{u}_n' A_n \mathbf{u}_n\right) = \left(E(u_i^4) - \sigma_u^4\right)\sum_i a_{ii}^2 - 2\sigma_u^4\sum_i a_{ii}^2 + \sigma_u^4\left[tr(A_n^2) + tr(A_n A_n')\right]$$

$$= (E(u_i^4) - 3\sigma_u^4)\sum_i a_{ii}^2 + \sigma_u^4\left(tr(A_n^2) + tr(A_n A_n')\right)$$

Let c_a be the bound on column sum and row of matrix A_n, so that it is also the bound on the absolute value of each element in A_n. Then $tr(A_n) \leq \sum_i c_a = nc_a = O(n)$. Also, $\sum_i a_{ii}^2 \leq c_a \sum_i |a_{ii}| \leq nc_a^2 = O(n)$. Similarly, since A_n is uniformly bounded in sum and row column, A_n^2 and A_n' and $A_n A_n'$ are also uniformly bounded in sum and row columns, therefore $tr(A_n^2) = O(n)$ and $tr(A_n A_n') = O(n)$ as well. We have:

$$E(\mathbf{u}_n' A_n \mathbf{u}_n) = O(n)$$
$$\text{Var}(\mathbf{u}_n' A_n \mathbf{u}_n) = O(n)$$

So that $\text{Var}\left(\frac{1}{n}\mathbf{u}_n' A_n \mathbf{u}_n\right) \to 0$ as $n \to \infty$. By Chebyshev's inequality we have last result in the lemma.

Lemma A.2.6 *Let $\mathbf{u}_n' = (u_1, \ldots, u_n)$ where u_i are iid with mean zero and finite variance σ_u^2 and $E(u_i^4) < \infty$. Let B_n be a $k \times n$ nonstochastic matrix in which elements are bounded in aboslute value by $c_b < \infty$, then, if $\lim_{n\to\infty} B_n B_n'$ exists and p.d.*

1. $\frac{1}{\sqrt{n}} B_n \mathbf{u}_n = O_p(1)$
2. $\frac{1}{\sqrt{n}} B_n \mathbf{u}_n \xrightarrow{d} N(0, \sigma_u^2 \lim_{n\to\infty} B_n B_n'/n)$
3. $\frac{1}{n} B_n \mathbf{u}_n = o_p(1)$

Kelejian and Prucha (1998).

A.2.2 Results for stochastic network structure

Lemma A.2.7 *Let A_n be $n \times n$ stochastic exogenous matrix that is uniformly bounded in row and column sum in n by c_a. For all i, $\sum_j Cov(a_{ii}, a_{jj}) \leq c_{aa}$. Let $\mathbf{u}_n' = (u_1, \ldots, u_n)$ where u_i are iid with mean zero and finite variance σ_u^2 and $E(u_i^4) < \infty$ and uncorrelated with elements of A_n, we have the following results*

1. $\frac{1}{n}\mathbf{u}_n' A_n \mathbf{u}_n - \frac{1}{n} E(\mathbf{u}_n' A_n \mathbf{u}_n) = o_p(1)$

From Lemma **A.2.5**, we can write

Mean:

$$E(\mathbf{u}_n' A_n \mathbf{u}_n) = E\left[E(\mathbf{u}_n' A_n \mathbf{u}_n \mid A_n)\right] = \sigma_u^2 E(tr(A_n)) \leq O(n)$$

since $E(tr(A_n)) = \sum_i E(a_{ii}) \leq \sum_i E(|a_{ii}|) \leq \sum_i c_a = nc_a$

Variance:

$$\begin{aligned}
\text{Var}\left(\mathbf{u}_n' A_n \mathbf{u}_n\right) &= E\left[\text{Var}\left(\mathbf{u}_n' A_n \mathbf{u}_n \mid A_n\right)\right] + \text{Var}\left[E(\mathbf{u}_n' A_n \mathbf{u}_n \mid A_n)\right] \\
\text{Var}\left(\mathbf{u}_n' A_n \mathbf{u}_n \mid A_n\right) &= \left(E(u_i^4) - \sigma_u^4\right)\sum_i a_{ii}^2 - 2\sigma_u^4 \sum_i a_{ii}^2 + \sigma_u^4\left[tr(A_n^2) + tr(A_n A_n')\right] \\
&= (E(u_i^4) - 3\sigma_u^4)\sum_i a_{ii}^2 + \sigma_u^4\left(tr(A_n^2) + tr(A_n A_n')\right)
\end{aligned}$$

$$\begin{aligned}
\rightarrow E\left[\text{Var}\left(\mathbf{u}_n' A_n \mathbf{u}_n \mid A_n\right)\right] &= O(n) \\
\text{Var}\left[E(\mathbf{u}_n' A_n \mathbf{u}_n \mid A_n)\right] &= \sigma_u^4 \text{Var}(tr(A_n))
\end{aligned}$$

we want to bound $\text{Var}(tr(A_n))$ by $O(n^{2-\alpha})$ for $\alpha > 0$

$$\text{Var}(tr(A_n)) = \text{Var}\left(\sum_i a_{ii}\right) = \sum_i \sum_j \text{Cov}(a_{ii}, a_{jj})$$

If we can assume that for all i, $\sum_j \text{Cov}(a_{ii}, a_{jj}) \leq c_{aa}$, then we can write $\text{Var}(tr(A_n)) \leq n c_{aa}$. As a result $\text{Var}\left(\mathbf{u}_n' A_n \mathbf{u}_n\right) = O(n)$ and we can obtain the result of the lemma by Chebyshev's inequality.[1]

[1]In our case of network, this assumption holds true trivially for for $A_n = G_n^{o*}$ as $w_{ii}^* = 0 \forall i$. However, for $A_n = S_n^{-1} = (I_n - \lambda G_n^{o*})^{-1}$, it is unclear. A sufficient condition is that if we have independent subnetworks so that G_n^{o*} and as a result S_n^{-1} can be written as block diagonal matrix and that each network is bounded in size. This means a_{ii} in block k is independent of a_{jj} in block $l \neq k$. For the case of a single large network, it is unclear how these conditions on stochastic matrices would translate to the structure of network.

Suppose we can break down n into ng independent subnetworks where $n = \sum_{ig=1}^{ng} n_{ig}$. This means

$$\text{Var}(tr(A_n)) = \sum_{ig=1}^{ng} \text{Var}\left(\sum_{i \in ig} a_{ii}\right) = \sum_{ig=1}^{ng} O(n_{ig}^2)$$

Now we want to impose conditions such that $\sum_{ig=1}^{ng} O(n_{ig}^2) = O(n^{2-\alpha}) = O\left(\left(\sum_{ig=1}^{ng} n_{ig}\right)^{2-\alpha}\right)$ as above. Denote $\max(n_{ig}) = \overline{n}_g$ and $\min(n_{ig}) = \underline{n}_g$, then

$$\text{Var}(tr(A_n)) = ng \times O\left(\max(n_{ig})^2\right) = O(ng \times \overline{n}_g^2)$$

For instance, our bounded independent subnetwork assumption means that $O(ng \times \underline{n}_g) = O\left(\left(\sum_{ig=1}^{ng} n_{ig}\right)^{2-\alpha}\right)$ for $\alpha = 1$, which means for the assumption to satisfy we must have zero growth rate of each subnetwork so that $O(\overline{n}_g^2) = O(\underline{n}_g)$.
For $0 < \alpha < 1$:

$$O\left(\left(ng \times \underline{n}_g\right)^{2-\alpha}\right) = O(ng \times ng^{1-\alpha}) O(\underline{n}_g^{2-\alpha})$$

Lemma A.2.8 *Let* $\mathbf{u}_n' = (u_1, \ldots, u_n)$ *where* u_i *are iid with mean zero and finite variance* σ_u^2 *and* $E(u_i^4) < \infty$*. Let* B_n *be a* $k \times n$ *stochastic matrix in which elements are bounded in aboslute value by* $c_b < \infty$*, then* $\frac{1}{n} B_n \mathbf{u}_n = o_p(1)$

Consider the variance of the ith element of $B_n \mathbf{u}_n$, since $E(u_i u_j) = 0$ for $i \neq j$ and u_i independent of elements of B_n:

$$\mathrm{Var}(\mathbf{b}_i \mathbf{u}_n) = \sum_i \mathrm{Var}(b_{ij} u_i) = \sigma_u^2 \sum_i \mathrm{Var}(b_{ij}) = O(n)$$

Lemma A.2.9 *If* A_n *is* $n \times n$ *uniformly bounded stochastic matrices in both row and column sums and* $\mathbf{x}_n$ *is* $n \times 1$ *stochastic column vectors of elements that are bounded in absolute value* ($|x_i| \leq c_x < \infty \forall i, n$)*. Then each element of* $|\mathbf{x}_n' A_n|$ *is bounded in absolute value*

$$\mathbf{x}_n' \mathbf{a}_i = \sum_j x_j a_{ij} \leq c_x \sum_j |a_{ij}| \leq c_x c_a$$

Lemma A.2.10 *If* $\mathbf{x}_n$ *is* $n \times 1$ *stochastic column vector of elements that are bounded in absolute value* ($|x_i| \leq c_x < \infty \forall i, n$)*. If* $\sum_j Cov(x_i^2, x_j^2) < c_{xx} < \infty$ *uniformly in* n*, then* $\frac{1}{n} \mathbf{x}_n' \mathbf{x}_n - \frac{1}{n} E(\mathbf{x}_n' \mathbf{x}_n) = o_p(1)$

$$E(\mathbf{x}_n' \mathbf{x}_n) = E\left(\sum_i x_i^2\right) = \sum_i E(x_i^2) \leq \sum_i E(|x_i|^2) \leq n c_x^2 = O(n)$$
$$\mathrm{Var}(\mathbf{x}_n' \mathbf{x}_n) = \mathrm{Var}\left(\sum_i x_i^2\right) = \sum_i \sum_j \mathrm{Cov}(x_i^2, x_j^2) \leq n c_{xx} = O(n)$$

A sufficient condition for $\sum_j \mathrm{Cov}(x_i^2, x_j^2) \leq c_{xx} < \infty$ is again the case of multiple independent networks, which is as $n \rightarrow \infty$ means number of network increases.

Lemma A.2.11 *Let* A_n *be* $n \times n$ *stochastic matrix whose row and column sums are uniformly bounded in absolute value. Let* $\mathbf{v}_n$ *be a random vector with zero means and is uncorrelated with elements of* A_n*. Furthermore, let* $Var(\mathbf{v}_n) = \Sigma_v$*, each element* $\sigma_{ij} \leq c_v < \infty$ *is absolutely bounded in* n *and* $E(v_i v_j v_l v_k) \leq c_{vv} < \infty$ *for any* i, j, l, k *is also absolutely bounded. Then* $\frac{1}{n^{1+\alpha}} \mathbf{v}_n' A_n \mathbf{v}_n - \frac{1}{n^{1+\alpha}} E(\mathbf{v}_n' A_n \mathbf{v}_n) = o_p(1)$ *for* $\alpha > 0$

so that another sufficient condition is $O(\overline{n}_g^2) = O(ng^{1-\alpha} \times \underline{n}_g^{2-\alpha})$. Further assume all subnetworks grow with the same rate n_g then we want $O(n_g^{\alpha}) = O(ng^{1-\alpha})$.

Mean:

$$E\left[\mathbf{v}_n' A_n \mathbf{v}_n\right] = \sum_j \sum_i E(a_{ij} v_i v_j) = \sum_j \sum_i E(a_{ij})\sigma_{ij}$$

$$\leq c_v \sum_j E\left(\sum_i a_{ij}\right) \leq c_v \sum_j E\left(\sum_i |a_{ij}|\right) \leq n c_v c_a = O(n)$$

Second moment:

$$(\mathbf{v}_n' A_n \mathbf{v}_n)^2 = \left(\sum_j \sum_i v_i v_j a_{ij}\right)\left(\sum_l \sum_k v_k v_l a_{kl}\right) = \sum_j \sum_i \sum_l \sum_k v_i v_j v_k v_l a_{ij} a_{kl}$$

$$E\left[(\mathbf{v}_n' A_n \mathbf{v}_n)^2\right] \leq c_{vv} E\left[\sum_j \sum_i \sum_l \sum_k |a_{ij}||a_{kl}|\right] \leq c_u c_a^2 \sum_j \sum_l 1 = n^2 c_u c_a^2$$

So that $\mathrm{Var}(\mathbf{v}_n' A_n \mathbf{v}_n) = O(n^2)$

Lemma A.2.12 *Let B_n be $k \times n$ stochastic matrix of which each element is absolutely bounded by c_b. Let $\mathbf{v}_n$ be a random vector with mean zeros and is uncorrelated with elements of B_n. Let $Var(\mathbf{v}_n) = \Sigma_v$, each element $\sigma_{ij} \leq c_v < \infty$ is absolutely bounded in n. Then $\frac{1}{n^{1+\alpha}} B_n \mathbf{v}_n = o_p(1)$ for $\alpha > 0$*

Consider the diagonal elements of the covariance atrix of $B_n \mathbf{v}_n$. Since v_i independent of elements of B_n:

$$\mathrm{Var}(\mathbf{b}_i \mathbf{v}_n) = \sum_j \sum_k \mathrm{Cov}(b_{ij} v_j, b_{ik} v_k) \leq c_b^2 \sum_j \sum_k \mathrm{Cov}(v_j, v_k) \leq c_b^2 c_v n^2 = O(n^2)$$

For off-diagonal elements:

$$\mathrm{Cov}(\mathbf{b}_i \mathbf{v}_n, \mathbf{b}_j \mathbf{v}_n) = \mathrm{Cov}\left(\sum_k b_{ik} v_k, \sum_l b_{jl} v_l\right)$$

$$\leq c_b^2 \mathrm{Cov}\left[\sum_k v_k, \sum_l v_l\right] = c_b^2 \sum_j \sum_k \mathrm{Cov}(v_j, v_k)$$

$$\leq c_b^2 c_v n^2 = O(n^2)$$

Lemma A.2.13 *Let A_n be $n \times n$ stochastic matrix whose row and column sums are uniformly bounded in absolute value. Let $\mathbf{v}_n$ be a random vector with mean zeros and is uncorrelated with elements of A_n. Let $Var(\mathbf{v}_n) = \Sigma_v$, each element $\sigma_{ij} \leq c_v < \infty$ is absolutely bounded in n. Let $\mathbf{u}_n$ be also a random vector where u_i are iid with mean zero and finite variance σ_u^2 and $E(u_i^4) < \infty$. $\mathbf{u}_n$ is uncorrelated with $\mathbf{v}_n$ and elements of A_n. Then $\frac{1}{n}\mathbf{v}_n' A_n \mathbf{u}_n - \frac{1}{n}E(\mathbf{v}_n' A_n \mathbf{u}_n) = o_p(1)$*

Mean:

$$E\left(\mathbf{v}_n' A_n \mathbf{u}_n\right) = E\left[\sum_j \sum_i v_i a_{ij} u_j\right] = 0$$

Variance:

$$(\mathbf{v}_n' A_n \mathbf{u}_n)^2 = \left(\sum_j \sum_i v_i u_j a_{ij}\right)\left(\sum_l \sum_k v_k u_l a_{kl}\right) = \sum_j \sum_i \sum_l \sum_k v_i u_j v_k u_l a_{ij} a_{kl}$$

$$E\left[(\mathbf{v}_n' A_n \mathbf{u}_n)^2\right] = \sum_i \sum_j \sum_k E(u_j^2 v_i v_k) E(a_{ij} a_{kj})$$

$$\leq \sigma_u^2 c_v E\left[\sum_i \sum_j \sum_k |a_{ij}||a_{kj}|\right] \leq \sigma_u^2 c_v c_a E\left[\sum_i \sum_j |a_{ij}|\right] \leq \sigma_u^2 c_v c_a^2 n = O(n)$$

A.3 Proof of theorem 1

Let $\boldsymbol{\theta} = (\delta, \gamma', \lambda, \boldsymbol{\beta}'^x, \rho)'$ be the parameters to estimate, in which $\boldsymbol{\theta}^d = (\delta, \gamma')'$ and $\boldsymbol{\theta}^y = (\lambda, \boldsymbol{\beta}'^x, \rho)'$ are the vectors of parameters in the first stage and in the second stage, respectively. Let the subscript 0 denote the true parameter values. To summarize, our model is:

$$
\begin{aligned}
p_i^* &= Pr\left(\underbrace{\mathbf{z}_i\gamma + \delta\mathbf{g}_i^*\mathbf{p}_n^*}_{r_i(\boldsymbol{\theta}^d)} + v_i > 0\right) \quad i = 1, \dots, n \\
d_i &= 1\left(v_i > -r_i(\boldsymbol{\theta}^d)\right) \\
\mathbf{y}_n^* &= S_n^{-1}(\lambda)\left[\mathbf{X}_n\boldsymbol{\beta}^x + \rho b_i(\boldsymbol{\theta}^d)\right] + M_n\boldsymbol{\epsilon}_n \\
\mathbf{y}_n &= D_n\mathbf{y}_n^* \\
b_i(\boldsymbol{\theta}^d) &= d_i\frac{\phi(-r_i(\boldsymbol{\theta}^d))}{1-\Phi(-r_i(\boldsymbol{\theta}^d))} - (1-d_i)\frac{\phi(-r_i(\boldsymbol{\theta}^d))}{\Phi(-r_i(\boldsymbol{\theta}^d))}
\end{aligned} \tag{A.3.1}
$$

where $M_n = I_n$ if we have incomplete information in the outcome stage, and $M_n = S_n^{-1}(\lambda)$ if we have complete information instead. With the frequentist method, we can estimate the model in two step. In order to obtain $\widehat{\boldsymbol{\theta}}^d$, we follow Xu (2018) or Lee et al. (2014). Then we can construct the bias term $\mathbf{b}_n(\widehat{\boldsymbol{\theta}}^d)$ and estimate the second stage by NLS, following Wang and Lee (2013).

Let $\mathbf{X}_n(\boldsymbol{\theta}^d) = [\boldsymbol{\iota}_n, \mathbf{X}_n, \mathbf{b}_n(\gamma, \delta)]$ of dimension $n \times k$:

$$
J_n^1\mathbf{y}_n = \underbrace{J_n^1 S_n^{-1}(\lambda_0)\mathbf{X}_n(\boldsymbol{\theta}_0^d)\boldsymbol{\beta}_0}_{h_n(\boldsymbol{\theta}^y,\boldsymbol{\theta}^d)} + J_n^1 M_n\boldsymbol{\epsilon}_n \tag{A.3.2}
$$

in which $M_n = I_n$ under incomplete information and $M_n = S_n^{-1}(\lambda_0)$ under complete information. Suppose we have a consistent estimator for parameters of the first stage, $\widehat{\boldsymbol{\theta}}^d \xrightarrow{p} \boldsymbol{\theta}_0^d$ and therefore has generated regressors $\mathbf{X}_n(\widehat{\boldsymbol{\theta}^d})$. The NLS criterion function:

$$
q_n(\boldsymbol{\theta}^y; \widehat{\boldsymbol{\theta}}^d) = [\mathbf{y}_n^1 - h_n(\boldsymbol{\theta}^y, \widehat{\boldsymbol{\theta}}^d)]'[\mathbf{y}_n^1 - h_n(\boldsymbol{\theta}^y, \widehat{\boldsymbol{\theta}}^d)] \tag{A.3.3}
$$

A.3.1 Consistency

We want to show first of all i) $\frac{1}{n}q_n(\boldsymbol{\theta}^y; \widehat{\boldsymbol{\theta}}^d) \xrightarrow{p} \frac{1}{n}E\left[q_n(\boldsymbol{\theta}^y; \boldsymbol{\theta}_0^d)\right]$ uniformly over Θ^y, and secondly ii) $\boldsymbol{\theta}_0^y$ uniquely minimizes $\frac{1}{n}E\left[q_n(\boldsymbol{\theta}^y; \boldsymbol{\theta}_0^d)\right]$.

Uniqueness

We have, using Lemma A.2.4:

$$\begin{aligned}
h_n(\boldsymbol{\theta}_0^y, \boldsymbol{\theta}_0^d) - h_n(\boldsymbol{\theta}^y, \boldsymbol{\theta}_0^d) &= J_n^1 \left[S_n^{-1}(\lambda_0)\mathbf{X}_n(\boldsymbol{\theta}_0^d)\boldsymbol{\beta}_0 - S_n^{-1}(\lambda)\mathbf{X}_n(\boldsymbol{\theta}_0^d)\boldsymbol{\beta}\right] \\
&= J_n^1 S_n^{-1}(\lambda)\left[\mathbf{X}_n(\boldsymbol{\theta}_0^d)(\boldsymbol{\beta}_0 - \boldsymbol{\beta}) + (\lambda_0 - \lambda)G_n^{o*}S_n^{-1}(\lambda_0)\mathbf{X}_n(\boldsymbol{\theta}_0^d)\boldsymbol{\beta}_0\right]
\end{aligned} \tag{A.3.4}$$

$$\mathbf{y}_n^1 - h_n(\boldsymbol{\theta}^y, \boldsymbol{\theta}_0^d) = J_n^1 S_n^{-1}(\lambda)\underbrace{\left[\mathbf{X}_n(\boldsymbol{\theta}_0^d)(\boldsymbol{\beta}_0 - \boldsymbol{\beta}) + (\lambda_0 - \lambda)G_n^{o*}S_n^{-1}(\lambda_0)\mathbf{X}_n(\boldsymbol{\theta}_0^d)\boldsymbol{\beta}_0\right]}_{\mathbf{d}_n(\boldsymbol{\theta}^y)} + J_n^1 M_n \boldsymbol{\epsilon}_n \tag{A.3.5}$$

We have J_n^1 uniformly bounded, $S_n(\lambda_0)$ and $S_n^{-1}(\lambda_0)$ uniformly bounded in n, and $S_n(\lambda)$ and $S_n^{-1}(\lambda)$ uniformly bounded in λ and in n. Each element of $\mathbf{d}_n(\boldsymbol{\theta}^y)$ uniformly bounded in n and λ due to assumption of uniformly bounded elements of $\mathbf{X}_n$ and compact parameter space.

$$\begin{aligned}
\frac{1}{n}E\left[q_n(\boldsymbol{\theta}^y; \boldsymbol{\theta}_0^d)\right] &= \frac{1}{n}E\left[(\mathbf{y}_n^1 - h_n(\boldsymbol{\theta}^y, \boldsymbol{\theta}_0^d))'(\mathbf{y}_n^1 - h_n(\boldsymbol{\theta}^y, \boldsymbol{\theta}_0^d))\right] \\
&= \frac{1}{n}\underbrace{E\left[\mathbf{d}_n(\boldsymbol{\theta}^y)'S_n'^{-1}(\lambda)J_n'^1 J_n^1 S_n^{-1}(\lambda)\mathbf{d}_n(\boldsymbol{\theta}^y)\right]}_{O(n) \text{ by Lemma A.2.9, Lemma A.2.10}} + \sigma_0^2\frac{1}{n}\underbrace{E\left[tr\left(M_n' J_n'^1 J_n^1 M_n\right]\right)}_{O(n)}
\end{aligned} \tag{A.3.6}$$

Let $\mathbf{m}_n = J_n^1 S_n^{-1}(\lambda)[\mathbf{X}_n(\boldsymbol{\theta}_0^d), G_n^{o*}S_n^{-1}(\lambda_0)\mathbf{X}_n(\boldsymbol{\theta}_0^d)\boldsymbol{\beta}_0]$. So that $\frac{1}{n}E\left[q_n(\boldsymbol{\theta}^y; \boldsymbol{\theta}_0^d)\right]$ obtains unique minimizer at $\boldsymbol{\theta}_0^y$ if we have full column rank of $E(\mathbf{m}_n'\mathbf{m}_n)$ for any value λ in Λ.

Uniform convergence

Consider:

$$\begin{aligned}
X_n(\boldsymbol{\theta}_0^d) - \mathbf{X}_n(\widehat{\boldsymbol{\theta}^d}) &= \left[\mathbf{0}_{n\times(k-1)} \quad (\mathbf{b}_n(\gamma_0, \delta_0) - \mathbf{b}_n(\hat{\gamma}, \hat{\delta}))\right] = \left[\mathbf{0}_{n\times(k-1)} \quad \boldsymbol{\xi}_n\right] \\
S_n^{-1}(\lambda)\mathbf{X}_n(\widehat{\boldsymbol{\theta}^d})\boldsymbol{\beta} &= S_n^{-1}(\lambda)\mathbf{X}_n(\boldsymbol{\theta}_0^d)\boldsymbol{\beta} - S_n^{-1}(\lambda)\boldsymbol{\xi}_n\rho
\end{aligned} \tag{A.3.7}$$

So that:

$$
\begin{aligned}
h_n(\boldsymbol{\theta}_0) - h_n(\boldsymbol{\theta}^y, \widehat{\boldsymbol{\theta}}^d) &= J_n^1 \left[S_n^{-1}(\lambda_0)\mathbf{X}_n(\boldsymbol{\theta}_0^d)\beta_0 - S_n^{-1}(\lambda)\mathbf{X}_n(\widehat{\boldsymbol{\theta}}^d)\boldsymbol{\beta} \right] \\
&= J_n^1 \left[S_n^{-1}(\lambda_0)\mathbf{X}_n(\boldsymbol{\theta}_0^d)\beta_0 - S_n^{-1}(\lambda)\mathbf{X}_n(\boldsymbol{\theta}_0^d)\boldsymbol{\beta} + S_n^{-1}(\lambda)\boldsymbol{\xi}_n\rho \right] \\
&= J_n^1 S_n^{-1}(\lambda)\mathbf{d}_n(\boldsymbol{\theta}^y) + J_n^1 S_n^{-1}(\lambda)\boldsymbol{\xi}_n\rho \qquad \text{(A.3.8)} \\
\mathbf{y}_n^1 - h_n(\boldsymbol{\theta}^y, \widehat{\boldsymbol{\theta}}^d) &= J_n^1 S_n^{-1}(\lambda)\mathbf{d}_n(\boldsymbol{\theta}^y) + J_n^1 S_n^{-1}(\lambda)\boldsymbol{\xi}_n\rho + J_n^1 M_n\boldsymbol{\epsilon}_n \qquad \text{(A.3.9)} \\
q_n(\boldsymbol{\theta}^y; \widehat{\boldsymbol{\theta}}^d) &= q_n(\boldsymbol{\theta}^y; \boldsymbol{\theta}_0^d) + \rho\boldsymbol{\xi}_n' S_n'^{-1}(\lambda)J_n'^1[\mathbf{d}_n(\boldsymbol{\theta}^y) + J_n^1 M_n\boldsymbol{\epsilon}_n] \\
&\quad + [\mathbf{d}_n(\boldsymbol{\theta}^y) + J_n^1 M_n\boldsymbol{\epsilon}_n]' J_n^1 S_n^{-1}(\lambda)\boldsymbol{\xi}_n\rho \\
&\quad + (\rho)^2\boldsymbol{\xi}_n' S_n'^{-1}(\lambda)J_n'^1 J_n^1 S_n^{-1}(\lambda)\boldsymbol{\xi}_n \\
\frac{1}{n}q_n(\boldsymbol{\theta}^y; \widehat{\boldsymbol{\theta}}^d) - \frac{1}{n}E\left[q_n(\boldsymbol{\theta}^y; \boldsymbol{\theta}_0^d)\right] &= \frac{1}{n}q_n(\boldsymbol{\theta}^y; \boldsymbol{\theta}_0^d) - \frac{1}{n}E\left[q_n(\boldsymbol{\theta}^y; \boldsymbol{\theta}_0^d)\right] \qquad \text{(A.3.10)} \\
&\quad + \frac{1}{n}\rho\boldsymbol{\xi}_n' S_n'^{-1}(\lambda)J_n'^1[\mathbf{d}_n(\boldsymbol{\theta}^y) + J_n^1 M_n\boldsymbol{\epsilon}_n] \qquad \text{(A.3.11)} \\
&\quad + \frac{1}{n}[\mathbf{d}_n(\boldsymbol{\theta}^y) + J_n^1 M_n\boldsymbol{\epsilon}_n]' J_n^1 S_n^{-1}(\lambda)\boldsymbol{\xi}_n\rho \qquad \text{(A.3.12)} \\
&\quad + \frac{1}{n}(\rho)^2\boldsymbol{\xi}_n' S_n'^{-1}(\lambda)J_n'^1 J_n^1 S_n^{-1}(\lambda)\boldsymbol{\xi}_n \qquad \text{(A.3.13)}
\end{aligned}
$$

We have the first term on (A.3.10):

$$
\begin{aligned}
&\frac{1}{n}q_n(\boldsymbol{\theta}^y; \boldsymbol{\theta}_0^d) - \frac{1}{n}E\left[q_n(\boldsymbol{\theta}^y; \boldsymbol{\theta}_0^d)\right] = \\
&\frac{1}{n}\mathbf{d}_n(\boldsymbol{\theta}^y)' S_n'^{-1}(\lambda)J_n'^1 J_n^1 S_n^{-1}(\lambda)\mathbf{d}_n(\boldsymbol{\theta}^y) - \frac{1}{n}E\left[\mathbf{d}_n(\boldsymbol{\theta}^y)' S_n'^{-1}(\lambda)J_n'^1 J_n^1 S_n^{-1}(\lambda)\mathbf{d}_n(\boldsymbol{\theta}^y)\right] \qquad \text{(A.3.14)} \\
&+ \frac{1}{n}\boldsymbol{\epsilon}_n' M_n' J_n'^1 J_n^1 S_n^{-1}(\lambda)\mathbf{d}_n(\boldsymbol{\theta}^y) + \frac{1}{n}\mathbf{d}_n(\boldsymbol{\theta}^y)' S_n'^{-1}(\lambda)J_n'^1 J_n^1 M_n\boldsymbol{\epsilon}_n \qquad \text{(A.3.15)} \\
&+ \frac{1}{n}\boldsymbol{\epsilon}_n' M_n' J_n'^1 J_n^1 M_n\boldsymbol{\epsilon}_n - \frac{1}{n}\sigma_0^2 E\left[tr\left(M_n' J_n'^1 J_n^1 M_n\right)\right] \qquad \text{(A.3.16)}
\end{aligned}
$$

Using results from Lemma **A.2.9** and Lemma **A.2.10**, then the term on (A.3.14) is $o_p(1)$. For terms on (A.3.15), we use results from Lemma **A.2.8**. For terms on (A.3.16), we apply Lemma **A.2.7**. The reason is we have ϵ_i are iid with zero means and independent of stochastic matrix $S_n^{-1}(\lambda)\ \forall\lambda$. For some B_n and A_n satisfying condition in those lemmas:

$$\begin{aligned}\frac{1}{n}\boldsymbol{\epsilon}_n' M_n' J_n'^1 J_n^1 S_n^{-1}(\lambda)\mathbf{d}_n(\boldsymbol{\theta}^y) &= \frac{1}{n}\boldsymbol{\epsilon}_n' B_n = o_p(1)\\ \frac{1}{n}\mathbf{d}_n(\boldsymbol{\theta}^y)' S_n'^{-1}(\lambda) J_n'^1 J_n^1 M_n\boldsymbol{\epsilon}_n &= \frac{1}{n}B_n\boldsymbol{\epsilon}_n = o_p(1)\\ \frac{1}{n}\boldsymbol{\epsilon}_n' M_n' J_n'^1 J_n^1 M_n\boldsymbol{\epsilon}_n - \frac{1}{n}\sigma_0^2 E\left(tr\left(M_n' J_n'^1 J_n^1 M_n\right)\right) &= \frac{1}{n}\boldsymbol{\epsilon}_n' A_n\boldsymbol{\epsilon}_n - \frac{1}{n}E\left[\boldsymbol{\epsilon}_n' A_n\boldsymbol{\epsilon}_n\right] = o_p(1)\end{aligned}$$

Therefore, we have $\frac{1}{n}q_n(\boldsymbol{\theta}^y;\boldsymbol{\theta}_0^d) - \frac{1}{n}E\left[q_n(\boldsymbol{\theta}^y;\boldsymbol{\theta}_0^d)\right] = o_p(1)$

We now deal with the rest of the terms related to $\boldsymbol{\xi}_n$ on (A.3.11), (A.3.12) and (A.3.13). $\boldsymbol{\xi}_n$ is the error related to the estimation of $\boldsymbol{\theta}_0^d$. Suppose we have a regular root n estimator for parameters in the first stage: $\sqrt{n}(\widehat{\boldsymbol{\theta}}^d - \boldsymbol{\theta}_0^d) = O_p(1) \xrightarrow{d} N(0, \Sigma_{\boldsymbol{\theta}_0^d})$. We have $b_i(\cdot) : \mathbb{R}^k \to \mathbb{R}$.[2] Suppose $b_i(\cdot)$ has continuous first derivatives and $\nabla_{\boldsymbol{\theta}'} b(\boldsymbol{\theta}_0^d) \neq 0$. By delta method, for each observation, $-\zeta_i = \sqrt{n}\xi_i = \sqrt{n}\left[b_i(\widehat{\boldsymbol{\theta}}^d) - b_i(\boldsymbol{\theta}_0^d)\right] = O_p(1) \xrightarrow{d} N(0, \nabla_{\boldsymbol{\theta}'} b_i(\boldsymbol{\theta}_0^d)\Sigma_{\boldsymbol{\theta}_0^d}\nabla_{\boldsymbol{\theta}'} b_i(\boldsymbol{\theta}_0^d)')$.

We want to construct $\widehat{\boldsymbol{\theta}}_i^d$ such that the sampling error $\boldsymbol{\xi}_n$ is (asymptotically) uncorrelated with $\mathbf{m}_n$. The sampling error will have non-diagonal covariance matrix Σ_ζ. We assume further that covariance of ζ_i and ζ_j is absolutely bounded in n and so is any fourth cross moment of $E(\zeta_i\zeta_j\zeta_k\zeta_l)$. Consider the following terms for some B_n and A_n satisfying conditions, using Lemma **A.2.11** and Lemma **A.2.12**:

$$\begin{aligned}\frac{1}{n}\rho\boldsymbol{\xi}_n' S_n'^{-1}(\lambda) J_n'^1\mathbf{d}_n(\boldsymbol{\theta}^y) &= \frac{1}{n^{3/2}}\boldsymbol{\zeta}_n' B_n = \frac{1}{\sqrt{n}}O_p(1) = o_p(1)\\ \frac{1}{n}\mathbf{d}_n(\boldsymbol{\theta}^y)' J_n^1 S_n^{-1}(\lambda)\boldsymbol{\xi}_n\rho &= \frac{1}{n^{3/2}}B_n\boldsymbol{\zeta}_n = \frac{1}{\sqrt{n}}O_p(1) = o_p(1)\\ \frac{1}{n}(\rho)^2\boldsymbol{\xi}_n' S_n'^{-1}(\lambda) J_n'^1 J_n^1 S_n^{-1}(\lambda)\boldsymbol{\xi}_n &= \frac{1}{n^2}\boldsymbol{\zeta}_n' A_n\boldsymbol{\zeta}_n\\ &= o_p(1) + \frac{1}{n^2}E(\boldsymbol{\zeta}_n' A_n\boldsymbol{\zeta}_n) = o_p(1) + \frac{1}{n^2}O(n) = o_p(1) + o(1)\end{aligned}$$

Finally, we deal with the remaining terms related to $\boldsymbol{\xi}_n$ and $\boldsymbol{\epsilon}_n$ on line (A.3.11) and (A.3.12). $\boldsymbol{\epsilon}_n$ is uncorrelated with $\boldsymbol{\zeta}_n$ so that we can use Lemma **A.2.13**. For A_n satisfying

[2] Note that each b_i is different due to different covariates

the condition in the lemma:

$$\frac{1}{n}\rho\boldsymbol{\xi}_n' S_n'^{-1}(\lambda)J_n'^1 J_n^1 S_n^{-1}(\lambda_0)\boldsymbol{\epsilon}_n = \frac{1}{n^{3/2}}\boldsymbol{\zeta}_n' A_n \boldsymbol{\epsilon}_n = o_p(1) + o(1)$$
$$\frac{1}{n}[J_n^1 S_n^{-1}(\lambda_0)\boldsymbol{\epsilon}_n]' J_n^1 S_n^{-1}(\lambda)\boldsymbol{\xi}_n \rho = \frac{1}{n^{3/2}}\boldsymbol{\epsilon}_n' A_n \boldsymbol{\zeta}_n = o_p(1) + o(1)$$

So that for each θ^y in Θ^y

$$\frac{1}{n}q_n(\boldsymbol{\theta}^y;\boldsymbol{\theta}_0^d) - \frac{1}{n}E\left[q_n(\boldsymbol{\theta}^y;\boldsymbol{\theta}_0^d)\right] \xrightarrow{p} 0$$

Stochastic Equicontinuity

What is left is the stochastic equicontinuity. We want to show, suppresing $\widehat{\boldsymbol{\theta}}^d$ inside q_n below:

$$\frac{1}{n}\left\|q_n(\boldsymbol{\theta}_1^y) - q_n(\boldsymbol{\theta}_2^y)\right\| \leq B_n|\boldsymbol{\theta}_1^y - \boldsymbol{\theta}_2^y|$$

for some $B_n = O_p(1)$. we have:

$$\begin{aligned}
\frac{\partial q_n(\boldsymbol{\theta}^y)}{\partial \boldsymbol{\theta}'^y} &= -2(\mathbf{y}_n^1 - h_n(\boldsymbol{\theta}^y))'\frac{\partial h_n(\boldsymbol{\theta}^y)}{\partial \boldsymbol{\theta}^y} \\
&= -2(\mathbf{y}_n^1 - h_n(\boldsymbol{\theta}^y))'\left[J_n^1 S_n^{-1}(\lambda)G_n^{o*}S_n^{-1}(\lambda)\mathbf{X}_n(\widehat{\boldsymbol{\theta}^d})\boldsymbol{\beta},\quad J_n^1 S_n^{-1}(\lambda)\mathbf{X}_n(\widehat{\boldsymbol{\theta}^d})\right] \\
&= -2\left[J_n^1 S_n^{-1}(\lambda)\mathbf{d}_n(\boldsymbol{\theta}^y) + J_n^1 S_n^{-1}(\lambda)\boldsymbol{\xi}_n\rho + J_n^1 M_n\boldsymbol{\epsilon}_n\right]' \times \\
&\left[J_n^1 S_n^{-1}(\lambda)G_n^{o*}S_n^{-1}(\lambda)\mathbf{X}_n(\boldsymbol{\theta}_0^d)\boldsymbol{\beta} - J_n^1 S_n^{-1}(\lambda)G_n^{o*}S_n^{-1}(\lambda)\boldsymbol{\xi}_n,\quad J_n^1 S_n^{-1}(\lambda)\mathbf{X}_n(\boldsymbol{\theta}_0^d) - J_n^1 S_n^{-1}(\lambda)\boldsymbol{\xi}_n\right]
\end{aligned}$$

Since we assume $S_n^{-1}(\lambda)$ is uniformly bounded in $\lambda \in \Lambda$, and we have the $n \times 1$ stochastic matrix $\mathbf{d}_n(\boldsymbol{\theta}^y) = \mathbf{X}_n(\boldsymbol{\theta}_0^d)(\boldsymbol{\beta}_0 - \boldsymbol{\beta}) + (\lambda_0 - \lambda)G_n^{o*}S_n^{-1}(\lambda_0)\mathbf{X}_n(\boldsymbol{\theta}_0^d)\boldsymbol{\beta}_0$ is linear in $\boldsymbol{\beta}$ and λ and uniformly bounded by some constant in n and in Θ. We have each element of the matrix $1 \times (k+1)$:

$$\frac{1}{n}\frac{\partial q_n(\boldsymbol{\theta}^y)}{\partial \boldsymbol{\theta}'^y}$$

is some linear and quadratic combinations of terms below, for stochastic matrix A_n and

B_n:

$$\frac{1}{n}\boldsymbol{\xi}_n' B_n = o_p(1); \quad \frac{1}{n}\boldsymbol{\epsilon}_n' B_n = o_p(1)$$
$$\frac{1}{n}\boldsymbol{\xi}_n' A_n \boldsymbol{\xi}_n = op(1); \quad \frac{1}{n}\boldsymbol{\epsilon}_n' A_n \boldsymbol{\xi}_n = o_p(1)$$
$$\frac{1}{n} B_n A_n B_n' = O_p(1) \quad \text{by Lemma A.2.9 Lemma A.2.10}$$

which means there exists B_n such that $\left\| \frac{1}{n} \frac{\partial q_n(\boldsymbol{\theta}^y)}{\partial \boldsymbol{\theta}'^y} \right\| \leq B_n = O_p(1)$.

Chapter B: Appendix for Chapter 3

B.1 Data

Table B.1: Science sample - Summary statistics for individuals

	Full		Science (1)		Science (2)	
	Mean	Sd	Mean	Sd	Mean	Sd
Own						
Take Sciences	0.88	0.33	0.88	0.33	0.88	0.32
GPA - Sciences	2.80	1.01	2.84	1.00	2.86	1.00
Peers						
Take Sciences	0.69	0.41	0.89	0.22	0.89	0.20
GPA - Sciences	2.21	1.33	2.89	0.70	2.90	0.68
Own						
Age	15.8	1.25	15.8	1.23	15.8	1.21
Male	0.50	0.50	0.49	0.50	0.49	0.50
Grade	10.4	1.11	10.4	1.11	10.4	1.10
Hispanic	0.18	0.39	0.16	0.36	0.082	0.27
White	0.61	0.49	0.67	0.47	0.74	0.44
Black	0.18	0.39	0.15	0.36	0.15	0.35
Asian	0.072	0.26	0.071	0.26	0.058	0.23
Other races	0.13	0.34	0.12	0.33	0.087	0.28
Years in school	2.53	1.43	2.58	1.43	2.46	1.35
Live w mother	0.92	0.26	0.93	0.25	0.94	0.24
Live w father	0.77	0.42	0.78	0.42	0.80	0.40
Health	2.11	0.93	2.09	0.91	2.09	0.90
Mom edu not know	0.071	0.26	0.061	0.24	0.051	0.22
Mom edu less HS	0.11	0.31	0.100	0.30	0.080	0.27
Mom edu HS	0.47	0.50	0.48	0.50	0.50	0.50

Continued on next page

	Full		Science (1)		Science (2)	
	Mean	Sd	Mean	Sd	Mean	Sd
Mom edu college	0.27	0.44	0.29	0.45	0.31	0.46
Dad edu not know	0.072	0.26	0.064	0.24	0.053	0.22
Dad edu less HS	0.089	0.29	0.083	0.28	0.074	0.26
Dad edu HS	0.34	0.47	0.35	0.48	0.37	0.48
Dad edu college	0.26	0.44	0.28	0.45	0.31	0.46
Mom job other	0.10	0.30	0.092	0.29	0.082	0.27
Mom job homemaker	0.17	0.38	0.17	0.37	0.16	0.37
Mom job professionals	0.30	0.46	0.31	0.46	0.33	0.47
Mom job office/sales	0.22	0.42	0.23	0.42	0.24	0.43
Mom job manual/farmer	0.12	0.32	0.12	0.32	0.11	0.32
Mom job military	0.0055	0.074	0.0054	0.073	0.0053	0.073
Dad job other	0.090	0.29	0.086	0.28	0.080	0.27
Dad job homemaker	0.0043	0.065	0.0036	0.060	0.0031	0.055
Dad job professionals	0.27	0.45	0.29	0.45	0.32	0.47
Dad job office/sales	0.054	0.23	0.057	0.23	0.059	0.24
Dad job manual/farmer	0.30	0.46	0.30	0.46	0.30	0.46
Dad job military	0.038	0.19	0.039	0.19	0.040	0.19
Household size	4.26	1.15	4.24	1.13	4.19	1.10
Peers						
Age	12.4	6.51	12.6	6.41	13.5	5.61
Male	0.37	0.33	0.37	0.34	0.40	0.33
Grade	8.17	4.34	8.34	4.30	8.94	3.78
Hispanic	0.13	0.27	0.11	0.26	0.067	0.18
White	0.51	0.45	0.55	0.45	0.64	0.43
Black	0.14	0.31	0.12	0.30	0.12	0.30
Asian	0.053	0.18	0.053	0.18	0.045	0.16
Other races	0.095	0.19	0.087	0.20	0.070	0.17

Continued on next page

	Full		Science (1)		Science (2)	
	Mean	Sd	Mean	Sd	Mean	Sd
Years in school	2.07	1.51	2.13	1.53	2.16	1.35
Live w mother	0.73	0.40	0.75	0.40	0.81	0.36
Live w father	0.62	0.39	0.64	0.40	0.70	0.36
Health	1.64	0.99	1.65	0.98	1.77	0.89
Mom edu not know	0.049	0.13	0.042	0.13	0.038	0.12
Mom edu less HS	0.078	0.18	0.070	0.18	0.061	0.16
Mom edu HS	0.38	0.33	0.39	0.35	0.43	0.33
Mom edu college	0.23	0.28	0.25	0.30	0.28	0.30
Dad edu not know	0.051	0.13	0.044	0.13	0.040	0.12
Dad edu less HS	0.068	0.16	0.063	0.16	0.060	0.15
Dad edu HS	0.27	0.29	0.29	0.31	0.32	0.30
Dad edu college	0.22	0.29	0.24	0.31	0.28	0.31
Mom job other	0.071	0.16	0.064	0.16	0.063	0.15
Mom job homemaker	0.13	0.21	0.13	0.22	0.14	0.21
Mom job professionals	0.25	0.28	0.27	0.30	0.30	0.29
Mom job office/sales	0.18	0.24	0.19	0.25	0.21	0.25
Mom job manual/farmer	0.090	0.18	0.089	0.19	0.090	0.18
Mom job military	0.0037	0.035	0.0036	0.037	0.0038	0.035
Dad job other	0.067	0.15	0.063	0.16	0.066	0.15
Dad job homemaker	0.0031	0.033	0.0028	0.035	0.0025	0.030
Dad job professionals	0.23	0.29	0.25	0.31	0.29	0.31
Dad job office/sales	0.045	0.12	0.048	0.13	0.052	0.13
Dad job manual/farmer	0.24	0.28	0.24	0.29	0.26	0.28
Dad job military	0.030	0.10	0.032	0.11	0.033	0.11
Household size	3.34	1.84	3.37	1.81	3.58	1.59
Observations	58787		42227		27408	

(1): After removing observations with any missing variables. (2): After dropping subnetworks not meeting coverage requirement.

Table B.2: Social Studies sample - Summary statistics for individuals

	Full		Science (1)		Science (2)	
	Mean	Sd	Mean	Sd	Mean	Sd
Own						
Take Social Studies	0.86	0.35	0.86	0.35	0.85	0.36
GPA - Social Studies	2.86	1.00	2.89	0.99	2.91	0.99
Peers						
Take Social Studies	0.68	0.41	0.87	0.24	0.85	0.25
GPA - Social Studies	2.26	1.36	2.95	0.72	2.96	0.71
Own						
Age	15.8	1.25	15.8	1.23	15.7	1.21
Male	0.50	0.50	0.49	0.50	0.49	0.50
Grade	10.4	1.11	10.4	1.11	10.4	1.10
Hispanic	0.18	0.39	0.15	0.36	0.081	0.27
White	0.61	0.49	0.66	0.47	0.74	0.44
Black	0.18	0.39	0.16	0.36	0.15	0.35
Asian	0.072	0.26	0.071	0.26	0.059	0.24
Other races	0.13	0.34	0.12	0.33	0.088	0.28
Years in school	2.53	1.43	2.58	1.43	2.46	1.35
Live w mother	0.92	0.26	0.93	0.25	0.94	0.24
Live w father	0.77	0.42	0.78	0.41	0.80	0.40
Health	2.11	0.93	2.09	0.91	2.09	0.90
Mom edu not know	0.071	0.26	0.062	0.24	0.051	0.22
Mom edu less HS	0.11	0.31	0.100	0.30	0.080	0.27
Mom edu HS	0.47	0.50	0.48	0.50	0.50	0.50
Mom edu college	0.27	0.44	0.29	0.45	0.31	0.46
Dad edu not know	0.072	0.26	0.064	0.24	0.053	0.22
Dad edu less HS	0.089	0.29	0.083	0.28	0.073	0.26
Dad edu HS	0.34	0.47	0.35	0.48	0.37	0.48

Continued on next page

	Full		Science (1)		Science (2)	
	Mean	Sd	Mean	Sd	Mean	Sd
Dad edu college	0.26	0.44	0.28	0.45	0.31	0.46
Mom job other	0.10	0.30	0.092	0.29	0.082	0.27
Mom job homemaker	0.17	0.38	0.17	0.38	0.16	0.37
Mom job professionals	0.30	0.46	0.31	0.46	0.33	0.47
Mom job office/sales	0.22	0.42	0.23	0.42	0.24	0.43
Mom job manual/farmer	0.12	0.32	0.12	0.32	0.11	0.32
Mom job military	0.0055	0.074	0.0053	0.073	0.0053	0.073
Dad job other	0.090	0.29	0.086	0.28	0.080	0.27
Dad job homemaker	0.0043	0.065	0.0037	0.061	0.0030	0.054
Dad job professionals	0.27	0.45	0.29	0.45	0.32	0.47
Dad job office/sales	0.054	0.23	0.058	0.23	0.059	0.24
Dad job manual/farmer	0.30	0.46	0.30	0.46	0.30	0.46
Dad job military	0.038	0.19	0.039	0.19	0.040	0.20
Household size	4.26	1.15	4.24	1.13	4.19	1.10
Peers						
Age	12.4	6.51	12.6	6.41	13.5	5.61
Male	0.37	0.33	0.37	0.34	0.40	0.33
Grade	8.17	4.34	8.34	4.30	8.94	3.78
Hispanic	0.13	0.27	0.11	0.26	0.067	0.18
White	0.51	0.45	0.55	0.45	0.64	0.43
Black	0.14	0.31	0.12	0.30	0.12	0.30
Asian	0.053	0.18	0.053	0.18	0.045	0.16
Other races	0.095	0.19	0.087	0.20	0.070	0.17
Years in school	2.07	1.51	2.14	1.53	2.16	1.35
Live w mother	0.73	0.40	0.75	0.40	0.81	0.36
Live w father	0.62	0.39	0.64	0.40	0.70	0.36
Health	1.64	0.99	1.65	0.98	1.77	0.89

Continued on next page

	Full		Science (1)		Science (2)	
	Mean	Sd	Mean	Sd	Mean	Sd
Mom edu not know	0.049	0.13	0.042	0.13	0.038	0.12
Mom edu less HS	0.078	0.18	0.071	0.18	0.062	0.16
Mom edu HS	0.38	0.33	0.39	0.35	0.43	0.33
Mom edu college	0.23	0.28	0.25	0.30	0.28	0.30
Dad edu not know	0.051	0.13	0.044	0.13	0.040	0.12
Dad edu less HS	0.068	0.16	0.063	0.16	0.060	0.15
Dad edu HS	0.27	0.29	0.29	0.31	0.32	0.30
Dad edu college	0.22	0.29	0.24	0.31	0.28	0.31
Mom job other	0.071	0.16	0.064	0.16	0.063	0.15
Mom job homemaker	0.13	0.21	0.13	0.22	0.14	0.21
Mom job professionals	0.25	0.28	0.27	0.30	0.30	0.29
Mom job office/sales	0.18	0.24	0.19	0.25	0.21	0.25
Mom job manual/farmer	0.090	0.18	0.089	0.19	0.091	0.18
Mom job military	0.0037	0.035	0.0036	0.037	0.0038	0.036
Dad job other	0.067	0.15	0.063	0.16	0.066	0.15
Dad job homemaker	0.0031	0.033	0.0029	0.036	0.0024	0.031
Dad job professionals	0.23	0.29	0.25	0.31	0.29	0.31
Dad job office/sales	0.045	0.12	0.049	0.13	0.053	0.13
Dad job manual/farmer	0.24	0.28	0.24	0.29	0.26	0.28
Dad job military	0.030	0.10	0.032	0.11	0.033	0.11
Household size	3.34	1.84	3.37	1.81	3.58	1.59
Observations	58787		42140		27392	

(1): After removing observations with any missing variables. (2): After dropping subnetworks not meeting coverage requirement.

Table B.3: Science sample - Summary statistics for subnetworks

	Science (1)				Science (2)			
	Mean	Sd	Min	Max	Mean	Sd	Min	Max
Coverage	0.41	0.21	0.0017	0.76	0.59	0.063	0.51	0.76
Selection rate	0.88	0.12	0.29	1	0.88	0.083	0.64	1.00
Subnetwork size	435.3	383.2	1	1675	622.9	359.0	145	1675
Observations	97				44			

(1): After removing observations with any missing variables. (2): After dropping subnetworks not meeting coverage requirement.

Table B.4: Social Studies sample - Summary statistics for subnetworks

	Social Studies (1)				Social Studies (2)			
	Mean	Sd	Min	Max	Mean	Sd	Min	Max
Coverage	0.41	0.21	0.0013	0.75	0.59	0.063	0.50	0.75
Selection rate	0.88	0.13	0.38	1	0.85	0.11	0.53	1
Subnetwork size	430	384.2	1	1666	622.5	361.7	143	1666
Observations	98				44			

(1): After removing observations with any missing variables considered. (2): After dropping subnetworks not meeting coverage requirement.

Table B.5: Summary of school size

	Frequency	Percent
Small (1-400 students)	3	6.82
Medium (401-1000 students)	22	50.00
Large (1001-4000 students)	19	43.18
N	44	

Table B.6: Summary of school type

	Frequency	Percent
Public	39	88.64
Private	5	11.36
N	44	

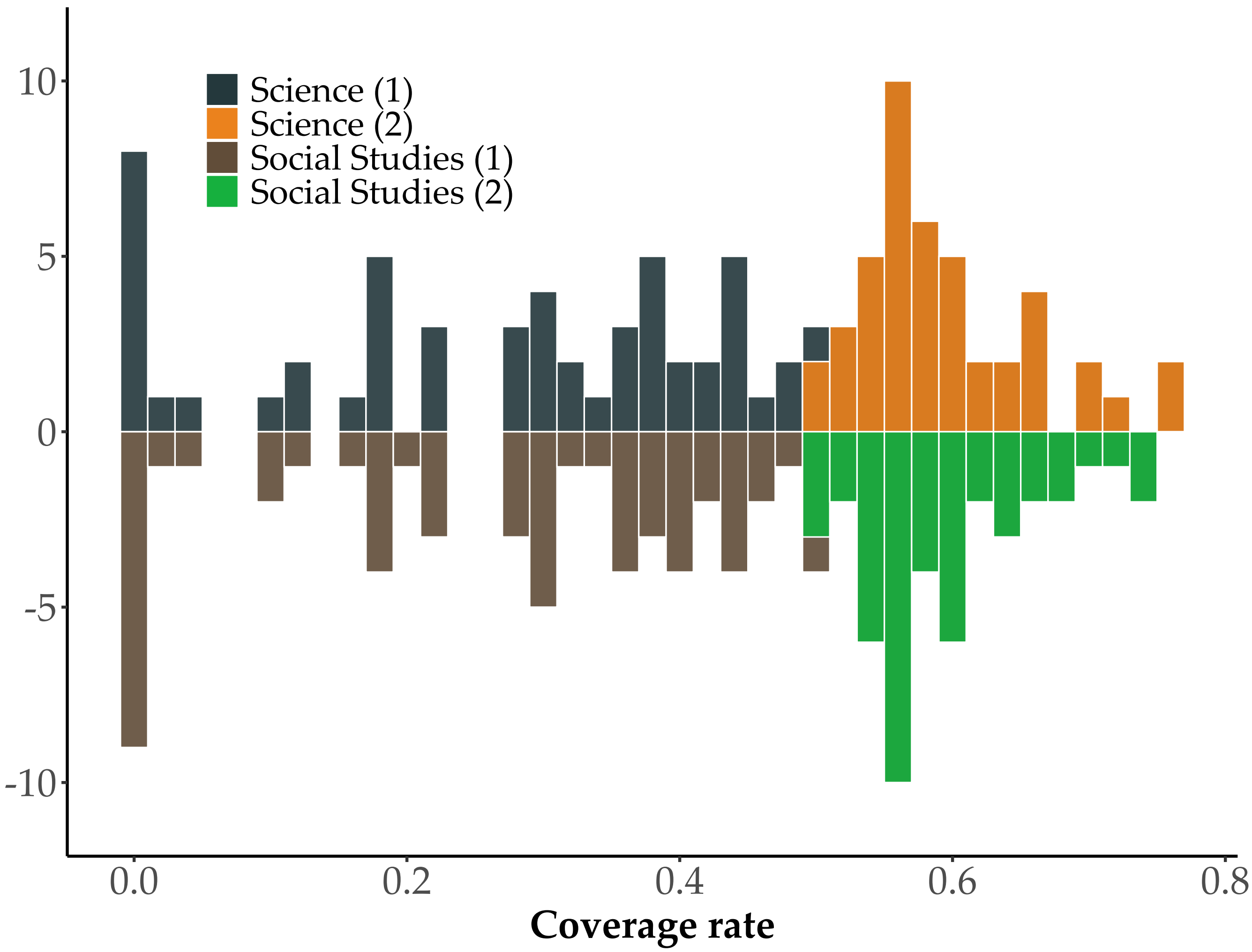
Table B.7: Summary of school region

	Frequency	Percent
West	5	11.36
Midwest	10	22.73
South	22	50.00
Northeast	7	15.91
N	44	

Table B.8: Summary of school location - Urbanicity

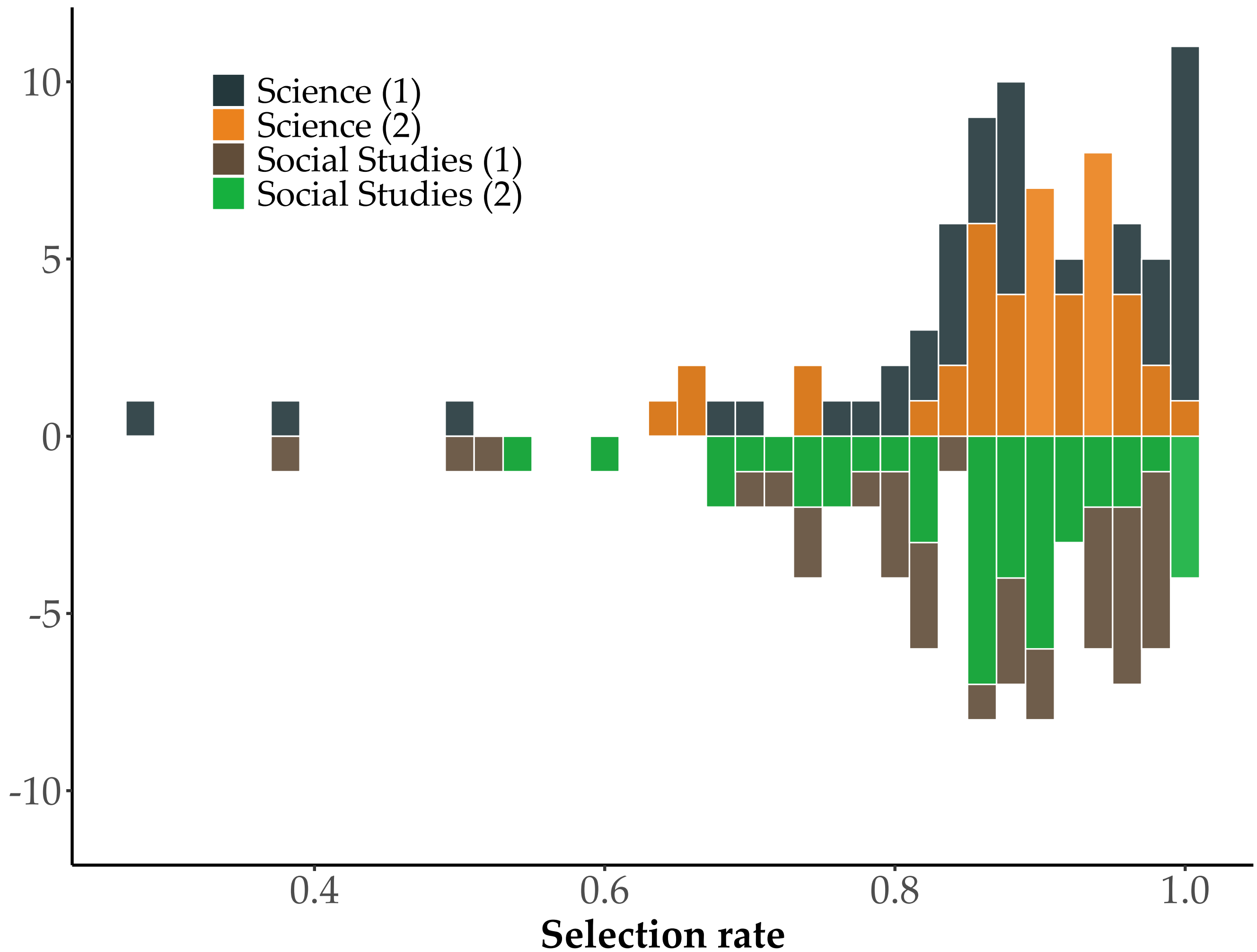
	Frequency	Percent
Urban	9	20.45
Suburban	29	65.91
Rural	6	13.64
N	44	

Figure B.1: Summary of coverage rate: Final sample size/School roster



Bin width = 0.02.
(1): After removing observations with any missing variables.
(2): After dropping subnetworks not meeting coverage requirement.

Figure B.2: Summary of average selection into subject of each school



Bin width = 0.02.
(1): After removing observations with any missing variables.
(2): After dropping subnetworks not meeting coverage requirement.

Table B.9: Subnetwork statistics

	Full	Science	Social Studies
Nodes - Roster	1053.4	1053.4	1053.4
Nodes - Sampled	811.9	622.9	622.5
Nodes - Isolated	27.4	24.3	24.7
(Proportion)	0.033	0.038	0.038
Coverage	0.78	0.59	0.59
Links	3617.8	2354.8	2354.1
Links - Reciprocal	1468.5	999.8	999
(Proportion)	0.41	0.42	0.42
Density (x100)	0.79	0.88	0.88
Degree	4.56	3.83	3.82
In-degree (sd)	3.51	3.06	3.05
In-degree (min)	0	0	0
In-degree (max)	21.5	18.4	18.4
Out-degree (sd)	2.90	2.47	2.47
Out-degree (min)	0	0	0
Out-degree (max)	9.89	9.73	9.70
Observations	44	44	44

The table reports the means of statistics of individual subnetworks. Full: The original network structure of the superset of subnetworks considered in all configurations.

B.2 Results - Details

B.2.1 Results without contextual peer effects

Table B.10: Science sample: Selection to take a course in Science

	Main (I)		Simple (IV)		Heckman (V)	
Endogenous peer effects (Δ)	0.144*	(0.056, 0.231)	-	-	-	-
Age	-0.090*	(-0.155, -0.026)	-0.096*	(-0.163, -0.030)	-0.096*	(-0.164, -0.029)
Male	0.025	(-0.038, 0.088)	0.014	(-0.044, 0.073)	0.014	(-0.044, 0.073)
Grade 10	0.279	(-0.008, 0.566)	0.284*	(0.028, 0.540)	0.284*	(0.028, 0.540)
Grade 11	-0.073	(-0.478, 0.332)	-0.072	(-0.427, 0.284)	-0.072	(-0.427, 0.284)
Grade 12	-0.429*	(-0.846, -0.013)	-0.434*	(-0.801, -0.067)	-0.434*	(-0.800, -0.067)
Hispanic	-0.068	(-0.182, 0.046)	-0.074	(-0.178, 0.030)	-0.074	(-0.178, 0.030)
Black	0.061	(-0.110, 0.232)	0.056	(-0.108, 0.219)	0.056	(-0.108, 0.219)
Asian	0.146	(-0.053, 0.344)	0.140	(-0.038, 0.319)	0.140	(-0.038, 0.319)
Other races	-0.123*	(-0.207, -0.039)	-0.122*	(-0.196, -0.048)	-0.122*	(-0.197, -0.047)
Years in school	-0.000	(-0.057, 0.057)	0.001	(-0.054, 0.056)	0.001	(-0.054, 0.056)
Live w mother	-0.051	(-0.167, 0.064)	-0.045	(-0.154, 0.065)	-0.044	(-0.154, 0.065)
Live w father	-0.088	(-0.409, 0.233)	-0.096	(-0.391, 0.200)	-0.096	(-0.392, 0.200)
Health	-0.053*	(-0.081, -0.025)	-0.055*	(-0.082, -0.028)	-0.055*	(-0.081, -0.028)
Mom edu not know	-0.067	(-0.171, 0.038)	-0.070	(-0.165, 0.024)	-0.070	(-0.166, 0.026)
Mom edu less HS	-0.018	(-0.104, 0.068)	-0.019	(-0.109, 0.071)	-0.019	(-0.109, 0.071)
Mom edu college	0.058	(-0.014, 0.130)	0.056	(-0.008, 0.121)	0.056	(-0.009, 0.122)
Dad edu not know	-0.066	(-0.172, 0.040)	-0.069	(-0.171, 0.033)	-0.069	(-0.172, 0.033)
Dad edu less HS	-0.114*	(-0.206, -0.023)	-0.114*	(-0.198, -0.029)	-0.114*	(-0.198, -0.030)
Dad edu college	0.107*	(0.035, 0.179)	0.108*	(0.036, 0.179)	0.107*	(0.038, 0.177)
Mom job other	0.109*	(0.023, 0.196)	0.107*	(0.023, 0.191)	0.107*	(0.023, 0.191)
Mom job professionals	0.132*	(0.049, 0.215)	0.134*	(0.050, 0.218)	0.134*	(0.052, 0.216)
Mom job office/sales	0.070	(-0.007, 0.146)	0.072	(-0.007, 0.151)	0.072	(-0.007, 0.150)
Mom job manual/farmer	0.033	(-0.053, 0.119)	0.033	(-0.051, 0.117)	0.033	(-0.053, 0.118)
Mom job military	-0.053	(-0.323, 0.217)	-0.052	(-0.285, 0.181)	-0.052	(-0.286, 0.182)
Dad job other	0.087	(-0.246, 0.421)	0.099	(-0.205, 0.403)	0.099	(-0.205, 0.403)
Dad job professionals	0.293	(-0.023, 0.609)	0.304*	(0.010, 0.598)	0.304*	(0.008, 0.600)
Dad job office/sales	0.281	(-0.034, 0.596)	0.293*	(0.007, 0.578)	0.293*	(0.006, 0.579)
Dad job manual/farmer	0.175	(-0.125, 0.476)	0.188	(-0.087, 0.464)	0.188	(-0.088, 0.465)
Dad job military	0.147	(-0.210, 0.504)	0.158	(-0.175, 0.492)	0.159	(-0.177, 0.494)
Household size	0.009	(-0.011, 0.029)	0.008	(-0.012, 0.027)	0.008	(-0.012, 0.027)
Small (1-400)	0.458	(-0.288, 1.204)	0.472	(-0.043, 0.987)	0.472	(-0.043, 0.986)
Large (1001-4000)	0.143	(-0.197, 0.483)	0.147	(-0.145, 0.438)	0.147	(-0.151, 0.444)

Continued on next page

Table B.10 – continued

	Main (I)		Simple (IV)		Heckman (V)	
Urban	0.132	(-0.251, 0.515)	0.129	(-0.164, 0.421)	0.128	(-0.175, 0.431)
Rural	0.214	(-0.104, 0.532)	0.224	(-0.050, 0.498)	0.224	(-0.051, 0.499)
West	-0.492	(-1.155, 0.172)	-0.512*	(-0.945, -0.079)	-0.512*	(-0.944, -0.080)
Midwest	-0.367	(-0.901, 0.167)	-0.375	(-0.753, 0.003)	-0.375	(-0.754, 0.004)
South	-0.179	(-0.710, 0.352)	-0.180	(-0.562, 0.202)	-0.180	(-0.563, 0.202)
Private	-0.029	(-0.802, 0.744)	-0.030	(-0.509, 0.450)	-0.030	(-0.511, 0.451)
Constant	2.636*	(1.498, 3.775)	2.841*	(1.721, 3.961)	2.842*	(1.708, 3.975)

[a] (I): MLE for strategic binary choice with incomplete information. (IV): Probit. (V): Heckit's 1st step. 95% confidence interval in parentheses, computed from 500 bootstraps for (I) and from clustered standard errors for (IV) and (V). * denotes CI does not include 0

Table B.11: Science sample: Determinants of Science GPAs - Point estimates

	Main (I)		(II)		(III)		Simple (IV)		Heckman (V)	
Endogenous peer effects (λ)	0.053*	(0.035, 0.070)	0.050*	(0.037, 0.063)	0.050*	(0.037, 0.064)	-	-	-	-
Age	-0.185*	(-0.228, -0.142)	-0.179*	(-0.216, -0.141)	-0.180*	(-0.217, -0.142)	-0.184*	(-0.224, -0.144)	-0.184*	(-0.224, -0.145)
Male	-0.118*	(-0.164, -0.073)	-0.120*	(-0.164, -0.076)	-0.120*	(-0.164, -0.076)	-0.134*	(-0.178, -0.090)	-0.134*	(-0.176, -0.092)
Grade 10	0.128*	(0.032, 0.224)	0.112*	(0.023, 0.200)	0.113*	(0.025, 0.201)	0.115*	(0.020, 0.210)	0.116*	(0.025, 0.207)
Grade 11	0.290*	(0.171, 0.409)	0.295*	(0.188, 0.401)	0.296*	(0.191, 0.402)	0.303*	(0.192, 0.414)	0.303*	(0.194, 0.412)
Grade 12	0.500*	(0.275, 0.725)	0.539*	(0.388, 0.690)	0.540*	(0.389, 0.690)	0.550*	(0.389, 0.712)	0.549*	(0.376, 0.721)
Hispanic	-0.135*	(-0.229, -0.040)	-0.129*	(-0.222, -0.035)	-0.129*	(-0.222, -0.036)	-0.133*	(-0.235, -0.031)	-0.133*	(-0.235, -0.031)
Black	-0.245*	(-0.370, -0.121)	-0.250*	(-0.375, -0.125)	-0.250*	(-0.374, -0.125)	-0.265*	(-0.391, -0.139)	-0.265*	(-0.388, -0.142)
Asian	0.214*	(0.117, 0.312)	0.205*	(0.119, 0.290)	0.205*	(0.120, 0.289)	0.198*	(0.117, 0.279)	0.198*	(0.120, 0.276)
Other races	-0.056	(-0.117, 0.004)	-0.047	(-0.095, 0.001)	-0.048	(-0.096, 0.000)	-0.048*	(-0.095, -0.001)	-0.049*	(-0.097, -0.000)
Years in school	0.032*	(0.002, 0.063)	0.032*	(0.002, 0.062)	0.032*	(0.002, 0.062)	0.034*	(0.005, 0.063)	0.034*	(0.005, 0.062)
Live w mother	0.183*	(0.118, 0.248)	0.186*	(0.124, 0.247)	0.185*	(0.123, 0.247)	0.192*	(0.127, 0.256)	0.191*	(0.128, 0.255)
Live w father	0.008	(-0.244, 0.259)	0.013	(-0.241, 0.267)	0.014	(-0.240, 0.269)	0.004	(-0.266, 0.274)	0.004	(-0.257, 0.265)
Health	-0.134*	(-0.157, -0.112)	-0.131*	(-0.148, -0.113)	-0.131*	(-0.148, -0.113)	-0.132*	(-0.151, -0.114)	-0.133*	(-0.151, -0.114)
Mom edu not know	0.014	(-0.051, 0.079)	0.019	(-0.043, 0.082)	0.020	(-0.043, 0.083)	0.016	(-0.049, 0.081)	0.016	(-0.047, 0.079)
Mom edu less HS	-0.136*	(-0.184, -0.089)	-0.135*	(-0.181, -0.089)	-0.136*	(-0.182, -0.090)	-0.141*	(-0.189, -0.093)	-0.141*	(-0.187, -0.094)
Mom edu college	0.119*	(0.078, 0.160)	0.116*	(0.080, 0.151)	0.116*	(0.080, 0.152)	0.114*	(0.076, 0.151)	0.114*	(0.075, 0.152)
Dad edu not know	-0.104*	(-0.175, -0.033)	-0.099*	(-0.167, -0.032)	-0.100*	(-0.167, -0.033)	-0.107*	(-0.177, -0.037)	-0.107*	(-0.175, -0.039)
Dad edu less HS	-0.072	(-0.152, 0.007)	-0.064	(-0.131, 0.004)	-0.064	(-0.131, 0.003)	-0.067	(-0.141, 0.007)	-0.067	(-0.144, 0.009)
Dad edu college	0.128*	(0.081, 0.175)	0.121*	(0.081, 0.162)	0.122*	(0.081, 0.162)	0.122*	(0.081, 0.163)	0.122*	(0.082, 0.162)
Mom job other	-0.107*	(-0.182, -0.032)	-0.115*	(-0.184, -0.046)	-0.115*	(-0.184, -0.045)	-0.117*	(-0.191, -0.043)	-0.117*	(-0.190, -0.043)
Mom job professionals	0.018	(-0.033, 0.068)	0.009	(-0.031, 0.049)	0.009	(-0.031, 0.049)	0.012	(-0.030, 0.053)	0.012	(-0.029, 0.054)
Mom job office/sales	-0.046*	(-0.092, -0.001)	-0.051*	(-0.093, -0.009)	-0.051*	(-0.093, -0.009)	-0.046*	(-0.088, -0.003)	-0.046*	(-0.087, -0.004)
Mom job manual/farmer	-0.088*	(-0.155, -0.021)	-0.090*	(-0.157, -0.024)	-0.090*	(-0.157, -0.024)	-0.090*	(-0.157, -0.023)	-0.090*	(-0.156, -0.025)
Mom job military	-0.073	(-0.294, 0.147)	-0.070	(-0.288, 0.149)	-0.072	(-0.293, 0.148)	-0.074	(-0.282, 0.134)	-0.074	(-0.275, 0.126)
Dad job other	-0.021	(-0.255, 0.214)	-0.026	(-0.264, 0.212)	-0.027	(-0.266, 0.211)	-0.014	(-0.272, 0.243)	-0.014	(-0.264, 0.235)
Dad job professionals	0.157	(-0.095, 0.409)	0.138	(-0.114, 0.390)	0.137	(-0.115, 0.389)	0.152	(-0.117, 0.421)	0.152	(-0.109, 0.414)
Dad job office/sales	0.104	(-0.144, 0.353)	0.086	(-0.162, 0.333)	0.085	(-0.164, 0.333)	0.099	(-0.164, 0.362)	0.100	(-0.156, 0.356)
Dad job manual/farmer	0.055	(-0.185, 0.295)	0.043	(-0.200, 0.286)	0.042	(-0.201, 0.285)	0.058	(-0.204, 0.319)	0.058	(-0.195, 0.311)

Continued on next page

Table B.11 – continued

	Main (I)		(II)		(III)		Simple (IV)		Heckman (V)	
Dad job military	0.082	(-0.188, 0.353)	0.072	(-0.203, 0.348)	0.070	(-0.206, 0.346)	0.087	(-0.206, 0.379)	0.087	(-0.196, 0.369)
Household size	-0.007	(-0.018, 0.005)	-0.007	(-0.018, 0.003)	-0.007	(-0.018, 0.003)	-0.009	(-0.020, 0.001)	-0.009	(-0.020, 0.001)
Small (1-400)	-0.026	(-0.422, 0.371)	-0.056	(-0.441, 0.330)	-0.054	(-0.438, 0.330)	-0.053	(-0.370, 0.265)	-0.052	(-0.368, 0.265)
Large (1001-4000)	-0.039	(-0.204, 0.126)	-0.050	(-0.209, 0.110)	-0.049	(-0.209, 0.110)	-0.053	(-0.192, 0.085)	-0.053	(-0.186, 0.081)
Urban	0.070	(-0.133, 0.273)	0.061	(-0.134, 0.257)	0.061	(-0.134, 0.256)	0.058	(-0.113, 0.230)	0.059	(-0.111, 0.228)
Rural	0.038	(-0.267, 0.343)	0.023	(-0.265, 0.312)	0.024	(-0.264, 0.311)	0.029	(-0.229, 0.287)	0.029	(-0.230, 0.289)
West	0.128	(-0.163, 0.419)	0.164	(-0.046, 0.375)	0.162	(-0.048, 0.371)	0.160*	(0.013, 0.308)	0.159	(-0.005, 0.323)
Midwest	0.080	(-0.161, 0.320)	0.105	(-0.090, 0.300)	0.104	(-0.090, 0.299)	0.109	(-0.024, 0.243)	0.108	(-0.030, 0.247)
South	0.029	(-0.201, 0.259)	0.041	(-0.163, 0.244)	0.041	(-0.163, 0.244)	0.040	(-0.119, 0.199)	0.040	(-0.120, 0.199)
Private	0.196	(-0.212, 0.604)	0.200	(-0.181, 0.580)	0.199	(-0.182, 0.579)	0.211	(-0.031, 0.453)	0.211	(-0.024, 0.446)
Constant	5.353*	(4.738, 5.969)	5.311*	(4.744, 5.878)	5.321*	(4.758, 5.884)	5.521*	(4.916, 6.125)	5.523*	(4.931, 6.114)
Selection	0.243	(-0.658, 1.145)	-	-	-	-	-	-	0.009	(-0.370, 0.387)

(I): NLS with bias correction and dynamic network. (II): NLS without bias correction and with dynamic network, (III): NLS without bias correction and with fixed network. (IV): Probit. (V): Heckit's 1st step. 95% confidence interval in parentheses, computed from 500 bootstraps for (I), (II), (III) and from clustered standard errors for (IV) and (V). * denotes CI does not include 0.

Table B.12: Science sample: Determinants of Science GPAs - Marginal effects

	Average own effects		Average total effects		OLS	
Age	-0.185	(-0.228, -0.142)	-0.194	(-0.240, -0.148)	-0.184	(-0.224, -0.144)
Male	-0.119	(-0.164, -0.073)	-0.124	(-0.172, -0.077)	-0.134	(-0.178, -0.090)
Grade 10	0.128	(0.032, 0.224)	0.134	(0.033, 0.235)	0.115	(0.020, 0.210)
Grade 11	0.290	(0.171, 0.409)	0.304	(0.180, 0.428)	0.303	(0.192, 0.414)
Grade 12	0.500	(0.275, 0.725)	0.523	(0.292, 0.755)	0.550	(0.389, 0.712)
Hispanic	-0.135	(-0.229, -0.040)	-0.141	(-0.240, -0.042)	-0.133	(-0.235, -0.031)
Black	-0.245	(-0.370, -0.121)	-0.257	(-0.386, -0.127)	-0.265	(-0.391, -0.139)
Asian	0.214	(0.117, 0.312)	0.224	(0.121, 0.328)	0.198	(0.117, 0.279)
Other races	-0.056	(-0.117, 0.004)	-0.059	(-0.122, 0.005)	-0.048	(-0.095, -0.001)
Years in school	0.032	(0.002, 0.063)	0.034	(0.002, 0.066)	0.034	(0.005, 0.063)
Live w mother	0.183	(0.118, 0.248)	0.191	(0.124, 0.259)	0.192	(0.127, 0.256)
Live w father	0.008	(-0.244, 0.259)	0.008	(-0.255, 0.271)	0.004	(-0.266, 0.274)
Health	-0.134	(-0.157, -0.112)	-0.141	(-0.166, -0.116)	-0.132	(-0.151, -0.114)
Mom edu not know	0.014	(-0.051, 0.079)	0.015	(-0.053, 0.083)	0.016	(-0.049, 0.081)
Mom edu less HS	-0.136	(-0.184, -0.089)	-0.143	(-0.192, -0.093)	-0.141	(-0.189, -0.093)
Mom edu college	0.119	(0.078, 0.160)	0.125	(0.081, 0.169)	0.114	(0.076, 0.151)
Dad edu not know	-0.104	(-0.175, -0.033)	-0.109	(-0.184, -0.034)	-0.107	(-0.177, -0.037)
Dad edu less HS	-0.072	(-0.152, 0.007)	-0.076	(-0.160, 0.008)	-0.067	(-0.141, 0.007)
Dad edu college	0.128	(0.081, 0.175)	0.134	(0.084, 0.184)	0.122	(0.081, 0.163)
Mom job other	-0.107	(-0.182, -0.032)	-0.112	(-0.190, -0.034)	-0.117	(-0.191, -0.043)
Mom job professionals	0.018	(-0.033, 0.069)	0.019	(-0.034, 0.072)	0.012	(-0.030, 0.053)
Mom job office/sales	-0.046	(-0.092, -0.001)	-0.049	(-0.096, -0.001)	-0.046	(-0.088, -0.003)
Mom job manual/farmer	-0.088	(-0.155, -0.021)	-0.092	(-0.162, -0.022)	-0.090	(-0.157, -0.023)
Mom job military	-0.074	(-0.294, 0.147)	-0.077	(-0.308, 0.154)	-0.074	(-0.282, 0.134)
Dad job other	-0.021	(-0.255, 0.214)	-0.021	(-0.267, 0.224)	-0.014	(-0.272, 0.243)
Dad job professionals	0.157	(-0.095, 0.410)	0.165	(-0.100, 0.429)	0.152	(-0.117, 0.421)
Dad job office/sales	0.105	(-0.144, 0.353)	0.109	(-0.151, 0.370)	0.099	(-0.164, 0.362)
Dad job manual/farmer	0.055	(-0.185, 0.295)	0.057	(-0.194, 0.308)	0.058	(-0.204, 0.319)
Dad job military	0.082	(-0.188, 0.353)	0.086	(-0.197, 0.369)	0.087	(-0.206, 0.379)
Household size	-0.007	(-0.018, 0.005)	-0.007	(-0.019, 0.005)	-0.009	(-0.020, 0.001)

Marginal effects for (I) in Table B.11. Marginal effects for students selecting the subject only. Marginal effects average within in each subnetwork then average over subnetworks. 95% confidence interval in parentheses, computed from 500 bootstraps.

Table B.13: Social Studies sample: Selection to take a course in Social Studies

	Main (I)		Simple (IV)		Heckman (V)	
Endogenous peer effects (δ)	0.070	(-0.009, 0.149)	-	-	-	-
Age	-0.001	(-0.054, 0.053)	-0.003	(-0.053, 0.046)	-0.004	(-0.054, 0.046)
Male	0.023	(-0.022, 0.069)	0.019	(-0.031, 0.069)	0.019	(-0.031, 0.069)
Grade 10	0.074	(-0.306, 0.454)	0.078	(-0.270, 0.425)	0.078	(-0.270, 0.425)
Grade 11	0.463*	(0.061, 0.864)	0.471*	(0.080, 0.861)	0.471*	(0.080, 0.862)
Grade 12	-0.107	(-0.520, 0.306)	-0.102	(-0.487, 0.283)	-0.101	(-0.487, 0.284)
Hispanic	-0.020	(-0.116, 0.075)	-0.021	(-0.110, 0.068)	-0.021	(-0.110, 0.068)
Black	0.068	(-0.089, 0.224)	0.066	(-0.073, 0.204)	0.065	(-0.073, 0.204)
Asian	-0.024	(-0.223, 0.175)	-0.027	(-0.203, 0.149)	-0.027	(-0.204, 0.149)
Other races	-0.108*	(-0.205, -0.010)	-0.107*	(-0.197, -0.017)	-0.107*	(-0.197, -0.017)
Years in school	0.036	(-0.035, 0.107)	0.037	(-0.035, 0.108)	0.037	(-0.035, 0.109)
Live w mother	-0.083	(-0.173, 0.006)	-0.079	(-0.171, 0.013)	-0.080	(-0.172, 0.013)
Live w father	-0.155	(-0.477, 0.168)	-0.155	(-0.430, 0.120)	-0.155	(-0.430, 0.120)
Health	-0.028*	(-0.050, -0.005)	-0.029*	(-0.052, -0.005)	-0.029*	(-0.052, -0.005)
Mom edu not know	-0.019	(-0.117, 0.079)	-0.022	(-0.119, 0.076)	-0.021	(-0.119, 0.076)
Mom edu less HS	0.029	(-0.040, 0.097)	0.028	(-0.038, 0.093)	0.027	(-0.038, 0.092)
Mom edu college	0.031	(-0.057, 0.119)	0.030	(-0.050, 0.111)	0.030	(-0.050, 0.111)
Dad edu not know	-0.071	(-0.180, 0.038)	-0.072	(-0.180, 0.036)	-0.073	(-0.180, 0.035)
Dad edu less HS	-0.048	(-0.137, 0.042)	-0.048	(-0.135, 0.039)	-0.048	(-0.135, 0.039)
Dad edu college	0.036	(-0.034, 0.106)	0.036	(-0.034, 0.106)	0.036	(-0.034, 0.105)
Mom job other	-0.022	(-0.097, 0.052)	-0.023	(-0.093, 0.046)	-0.024	(-0.093, 0.046)
Mom job professionals	0.062	(-0.007, 0.132)	0.062	(-0.007, 0.131)	0.062	(-0.007, 0.131)
Mom job office/sales	0.050	(-0.009, 0.109)	0.051	(-0.009, 0.110)	0.050	(-0.009, 0.110)
Mom job manual/farmer	0.032	(-0.048, 0.113)	0.031	(-0.049, 0.110)	0.030	(-0.049, 0.110)
Mom job military	0.149	(-0.185, 0.482)	0.145	(-0.161, 0.451)	0.146	(-0.161, 0.453)
Dad job other	0.126	(-0.225, 0.477)	0.128	(-0.176, 0.431)	0.128	(-0.175, 0.432)
Dad job professionals	0.235	(-0.110, 0.581)	0.237	(-0.066, 0.540)	0.237	(-0.065, 0.540)
Dad job office/sales	0.286	(-0.064, 0.637)	0.288	(-0.025, 0.602)	0.288	(-0.025, 0.602)
Dad job manual/farmer	0.210	(-0.115, 0.535)	0.213	(-0.061, 0.487)	0.213	(-0.061, 0.487)
Dad job military	0.231	(-0.143, 0.604)	0.231	(-0.090, 0.551)	0.231	(-0.090, 0.552)
Household size	0.010	(-0.014, 0.034)	0.010	(-0.013, 0.033)	0.010	(-0.013, 0.033)
Small (1-400)	-0.155	(-0.651, 0.342)	-0.157	(-0.490, 0.176)	-0.157	(-0.492, 0.178)
Large (1001-4000)	0.103	(-0.240, 0.447)	0.104	(-0.168, 0.376)	0.104	(-0.168, 0.376)
Urban	0.134	(-0.199, 0.468)	0.134	(-0.120, 0.387)	0.134	(-0.120, 0.389)
Rural	0.375	(-0.085, 0.835)	0.382*	(0.046, 0.718)	0.383*	(0.045, 0.720)
West	-0.379	(-2.037, 1.279)	-0.388	(-0.889, 0.112)	-0.388	(-0.889, 0.113)
Midwest	-0.524	(-2.175, 1.128)	-0.531	(-1.102, 0.040)	-0.531	(-1.103, 0.041)

Continued on next page

Table B.13 – continued

	Main (I)		Simple (IV)		Heckman (V)	
South	-0.217	(-1.835, 1.401)	-0.220	(-0.693, 0.253)	-0.219	(-0.693, 0.254)
Private	-0.084	(-1.444, 1.275)	-0.087	(-0.461, 0.287)	-0.087	(-0.462, 0.288)
Constant	0.978	(-0.854, 2.811)	1.071*	(0.059, 2.083)	1.075*	(0.062, 2.087)

(I): MLE for strategic binary choice with incomplete information. (IV): Probit. (V): Heckit's 1st step. 95% confidence interval in parentheses, computed from 500 bootstraps for (I) and from clustered standard errors for (IV) and (V). * denotes CI does not include 0.

Table B.14: Social Studies sample: Determinants of Social Studies GPAs - Point estimates

	Main (I)		(II)		(III)		Simple (IV)		Heckman (V)	
Endogenous peer effects (λ)	0.061*	(0.041, 0.081)	0.053*	(0.036, 0.071)	0.062*	(0.044, 0.079)	-	-	-	-
Age	-0.187*	(-0.238, -0.136)	-0.187*	(-0.233, -0.141)	-0.186*	(-0.233, -0.140)	-0.193*	(-0.243, -0.144)	-0.193*	(-0.241, -0.145)
Male	-0.076*	(-0.124, -0.028)	-0.086*	(-0.126, -0.045)	-0.083*	(-0.124, -0.043)	-0.100*	(-0.144, -0.057)	-0.100*	(-0.143, -0.057)
Grade 10	0.150	(-0.023, 0.324)	0.117*	(0.014, 0.220)	0.118*	(0.015, 0.220)	0.121*	(0.011, 0.231)	0.122*	(0.016, 0.228)
Grade 11	0.492*	(0.231, 0.753)	0.328*	(0.210, 0.446)	0.329*	(0.211, 0.446)	0.336*	(0.209, 0.463)	0.342*	(0.221, 0.462)
Grade 12	0.586*	(0.368, 0.805)	0.631*	(0.454, 0.808)	0.629*	(0.453, 0.806)	0.645*	(0.457, 0.834)	0.644*	(0.462, 0.826)
Hispanic	-0.164*	(-0.240, -0.089)	-0.156*	(-0.234, -0.078)	-0.155*	(-0.233, -0.077)	-0.161*	(-0.242, -0.080)	-0.162*	(-0.240, -0.084)
Black	-0.234*	(-0.335, -0.133)	-0.264*	(-0.355, -0.172)	-0.261*	(-0.353, -0.170)	-0.281*	(-0.369, -0.192)	-0.280*	(-0.365, -0.194)
Asian	0.199*	(0.088, 0.311)	0.209*	(0.112, 0.306)	0.210*	(0.113, 0.307)	0.201*	(0.110, 0.293)	0.201*	(0.112, 0.289)
Other races	-0.112*	(-0.201, -0.024)	-0.066*	(-0.127, -0.005)	-0.067*	(-0.128, -0.007)	-0.068*	(-0.127, -0.010)	-0.070*	(-0.126, -0.014)
Years in school	0.032	(-0.010, 0.075)	0.018	(-0.013, 0.049)	0.018	(-0.013, 0.049)	0.020	(-0.012, 0.052)	0.021	(-0.011, 0.052)
Live w mother	0.137*	(0.057, 0.217)	0.169*	(0.100, 0.238)	0.167*	(0.099, 0.236)	0.176*	(0.104, 0.248)	0.175*	(0.105, 0.245)
Live w father	-0.040	(-0.309, 0.228)	0.028	(-0.219, 0.274)	0.029	(-0.216, 0.274)	0.007	(-0.241, 0.254)	0.004	(-0.234, 0.242)
Health	-0.148*	(-0.170, -0.125)	-0.137*	(-0.152, -0.122)	-0.137*	(-0.152, -0.122)	-0.139*	(-0.155, -0.123)	-0.139*	(-0.155, -0.124)
Mom edu not know	-0.042	(-0.111, 0.028)	-0.033	(-0.085, 0.019)	-0.033	(-0.085, 0.019)	-0.038	(-0.088, 0.012)	-0.038	(-0.087, 0.011)
Mom edu less HS	-0.136*	(-0.191, -0.081)	-0.147*	(-0.197, -0.097)	-0.147*	(-0.197, -0.097)	-0.152*	(-0.208, -0.097)	-0.152*	(-0.206, -0.098)
Mom edu college	0.153*	(0.109, 0.196)	0.140*	(0.108, 0.172)	0.140*	(0.108, 0.173)	0.139*	(0.108, 0.169)	0.139*	(0.109, 0.170)
Dad edu not know	-0.183*	(-0.266, -0.101)	-0.152*	(-0.217, -0.088)	-0.150*	(-0.214, -0.086)	-0.158*	(-0.222, -0.094)	-0.159*	(-0.221, -0.098)
Dad edu less HS	-0.112*	(-0.182, -0.042)	-0.093*	(-0.156, -0.030)	-0.092*	(-0.155, -0.029)	-0.093*	(-0.161, -0.026)	-0.094*	(-0.159, -0.028)
Dad edu college	0.189*	(0.139, 0.240)	0.176*	(0.134, 0.218)	0.176*	(0.134, 0.219)	0.178*	(0.133, 0.223)	0.178*	(0.135, 0.222)
Mom job other	-0.182*	(-0.266, -0.098)	-0.171*	(-0.243, -0.098)	-0.171*	(-0.244, -0.099)	-0.174*	(-0.253, -0.094)	-0.174*	(-0.251, -0.097)
Mom job professionals	0.030	(-0.027, 0.087)	0.008	(-0.042, 0.057)	0.008	(-0.041, 0.058)	0.011	(-0.041, 0.063)	0.012	(-0.038, 0.062)
Mom job office/sales	-0.031	(-0.090, 0.028)	-0.049	(-0.102, 0.004)	-0.049	(-0.102, 0.004)	-0.043	(-0.100, 0.013)	-0.043	(-0.098, 0.013)
Mom job manual/farmer	-0.082*	(-0.150, -0.014)	-0.093*	(-0.155, -0.030)	-0.092*	(-0.154, -0.030)	-0.091*	(-0.160, -0.022)	-0.091*	(-0.158, -0.024)
Mom job military	-0.076	(-0.259, 0.107)	-0.131	(-0.275, 0.014)	-0.133	(-0.278, 0.013)	-0.127	(-0.270, 0.016)	-0.125	(-0.264, 0.014)
Dad job other	-0.017	(-0.288, 0.255)	-0.074	(-0.313, 0.165)	-0.076	(-0.314, 0.162)	-0.051	(-0.293, 0.192)	-0.049	(-0.282, 0.185)
Dad job professionals	0.203	(-0.084, 0.489)	0.103	(-0.159, 0.364)	0.101	(-0.159, 0.362)	0.129	(-0.136, 0.394)	0.132	(-0.121, 0.386)
Dad job office/sales	0.239	(-0.055, 0.532)	0.119	(-0.145, 0.384)	0.117	(-0.146, 0.380)	0.143	(-0.122, 0.409)	0.148	(-0.106, 0.402)
Dad job manual/farmer	0.083	(-0.191, 0.358)	-0.007	(-0.254, 0.239)	-0.009	(-0.255, 0.236)	0.018	(-0.232, 0.269)	0.022	(-0.218, 0.261)
Dad job military	0.186	(-0.097, 0.468)	0.091	(-0.166, 0.347)	0.087	(-0.167, 0.342)	0.116	(-0.147, 0.379)	0.120	(-0.132, 0.371)

Continued on next page

Table B.14 – continued

	Main (I)		(II)		(III)		Simple (IV)		Heckman (V)	
Household size	-0.002	(-0.020, 0.016)	-0.006	(-0.022, 0.009)	-0.006	(-0.021, 0.009)	-0.008	(-0.023, 0.007)	-0.008	(-0.023, 0.007)
Small (1-400)	-0.065	(-0.544, 0.414)	0.011	(-0.412, 0.433)	0.011	(-0.408, 0.431)	0.015	(-0.336, 0.366)	0.012	(-0.330, 0.354)
Large (1001-4000)	-0.009	(-0.176, 0.158)	-0.050	(-0.199, 0.099)	-0.050	(-0.197, 0.098)	-0.053	(-0.187, 0.080)	-0.052	(-0.180, 0.076)
Urban	0.023	(-0.187, 0.233)	-0.031	(-0.199, 0.136)	-0.030	(-0.196, 0.136)	-0.039	(-0.168, 0.091)	-0.037	(-0.164, 0.091)
Rural	0.239	(-0.036, 0.513)	0.103	(-0.107, 0.313)	0.103	(-0.105, 0.311)	0.113	(-0.071, 0.297)	0.118	(-0.060, 0.296)
West	-0.005	(-0.277, 0.266)	0.143	(-0.060, 0.347)	0.142	(-0.060, 0.344)	0.137	(-0.020, 0.294)	0.132	(-0.019, 0.282)
Midwest	-0.186	(-0.583, 0.212)	0.023	(-0.189, 0.235)	0.021	(-0.190, 0.232)	0.021	(-0.152, 0.194)	0.014	(-0.168, 0.195)
South	-0.024	(-0.278, 0.230)	0.058	(-0.147, 0.263)	0.057	(-0.146, 0.261)	0.059	(-0.103, 0.220)	0.056	(-0.101, 0.213)
Private	0.027	(-0.383, 0.438)	0.061	(-0.277, 0.398)	0.058	(-0.278, 0.393)	0.068	(-0.152, 0.289)	0.067	(-0.149, 0.283)
Constant	5.200*	(4.244, 6.156)	5.543*	(4.819, 6.266)	5.510*	(4.793, 6.227)	5.777*	(5.024, 6.531)	5.766*	(5.039, 6.494)
Selection	1.130	(-0.401, 2.662)	-	-	-	-	-	-	0.040	(-0.067, 0.148)

(I): NLS with bias correction and dynamic network. (II): NLS without bias correction and with dynamic network, (III): NLS without bias correction and with fixed network. (IV): Probit. (V): Heckit's 1st step. 95% confidence interval in parentheses, computed from 500 bootstraps for (I), (II), (III) and from clustered standard errors for (IV) and (V). * denotes CI does not include 0.

Table B.15: Social Studies sample: Determinants of Social Studies GPAs - Marginal effects

	Average own effects		Average total effects		OLS	
Age	-0.187	(-0.238, -0.136)	-0.197	(-0.251, -0.144)	-0.193	(-0.243, -0.144)
Male	-0.076	(-0.124, -0.028)	-0.080	(-0.131, -0.030)	-0.100	(-0.144, -0.057)
Grade 10	0.150	(-0.023, 0.324)	0.159	(-0.025, 0.342)	0.121	(0.011, 0.231)
Grade 11	0.492	(0.231, 0.754)	0.519	(0.243, 0.795)	0.336	(0.209, 0.463)
Grade 12	0.587	(0.368, 0.805)	0.618	(0.389, 0.847)	0.645	(0.457, 0.834)
Hispanic	-0.164	(-0.240, -0.089)	-0.173	(-0.252, -0.094)	-0.161	(-0.242, -0.080)
Black	-0.234	(-0.335, -0.133)	-0.246	(-0.352, -0.141)	-0.281	(-0.369, -0.192)
Asian	0.199	(0.088, 0.311)	0.210	(0.092, 0.328)	0.201	(0.110, 0.293)
Other races	-0.112	(-0.201, -0.024)	-0.118	(-0.212, -0.025)	-0.068	(-0.127, -0.010)
Years in school	0.032	(-0.010, 0.075)	0.034	(-0.011, 0.079)	0.020	(-0.012, 0.052)
Live w mother	0.137	(0.057, 0.217)	0.144	(0.060, 0.229)	0.176	(0.104, 0.248)
Live w father	-0.040	(-0.309, 0.228)	-0.042	(-0.325, 0.240)	0.007	(-0.241, 0.254)
Health	-0.148	(-0.170, -0.125)	-0.156	(-0.181, -0.131)	-0.139	(-0.155, -0.123)
Mom edu not know	-0.042	(-0.111, 0.028)	-0.044	(-0.117, 0.030)	-0.038	(-0.088, 0.012)
Mom edu less HS	-0.136	(-0.191, -0.081)	-0.143	(-0.201, -0.085)	-0.152	(-0.208, -0.097)
Mom edu college	0.153	(0.109, 0.196)	0.161	(0.114, 0.208)	0.139	(0.108, 0.169)
Dad edu not know	-0.184	(-0.266, -0.101)	-0.193	(-0.281, -0.105)	-0.158	(-0.222, -0.094)
Dad edu less HS	-0.112	(-0.182, -0.042)	-0.118	(-0.192, -0.044)	-0.093	(-0.161, -0.026)
Dad edu college	0.189	(0.139, 0.240)	0.200	(0.145, 0.254)	0.178	(0.133, 0.223)
Mom job other	-0.182	(-0.266, -0.098)	-0.192	(-0.281, -0.104)	-0.174	(-0.253, -0.094)
Mom job professionals	0.030	(-0.027, 0.087)	0.032	(-0.029, 0.092)	0.011	(-0.041, 0.063)
Mom job office/sales	-0.031	(-0.090, 0.028)	-0.033	(-0.095, 0.030)	-0.043	(-0.100, 0.013)
Mom job manual/farmer	-0.082	(-0.150, -0.014)	-0.086	(-0.158, -0.014)	-0.091	(-0.160, -0.022)
Mom job military	-0.076	(-0.259, 0.107)	-0.080	(-0.272, 0.112)	-0.127	(-0.270, 0.016)
Dad job other	-0.017	(-0.288, 0.255)	-0.018	(-0.303, 0.268)	-0.051	(-0.293, 0.192)
Dad job professionals	0.203	(-0.084, 0.489)	0.214	(-0.089, 0.517)	0.129	(-0.136, 0.394)
Dad job office/sales	0.239	(-0.055, 0.532)	0.252	(-0.058, 0.562)	0.143	(-0.122, 0.409)
Dad job manual/farmer	0.083	(-0.191, 0.358)	0.088	(-0.201, 0.377)	0.018	(-0.232, 0.269)
Dad job military	0.186	(-0.097, 0.468)	0.196	(-0.102, 0.494)	0.116	(-0.147, 0.379)
Household size	-0.002	(-0.020, 0.016)	-0.002	(-0.021, 0.017)	-0.008	(-0.023, 0.007)

Marginal effects for (I) in Table B.14. Marginal effects for students selecting the subject only. Marginal effects average within in each subnetwork then average over subnetworks. 95% confidence interval in parentheses, computed from 500 bootstraps.

B.2.2 Results with contextual peer effects

Table B.16: Science sample: Selection to take a course in Science with contextual peer effects

	Main (I)	
Endogenous peer effects(δ)	0.592	(-0.726, 1.909)
Own effects		
Age	-0.083*	(-0.149, -0.017)
Male	0.010	(-0.045, 0.065)
Grade 10	0.533*	(0.217, 0.849)
Grade 11	0.067	(-0.365, 0.500)
Grade 12	-0.387	(-0.801, 0.027)
Hispanic	-0.058	(-0.148, 0.033)
Black	0.146*	(0.013, 0.279)
Asian	0.299*	(0.164, 0.434)
Other races	-0.099*	(-0.173, -0.025)
Years in school	-0.010	(-0.050, 0.030)
Live w mother	-0.072	(-0.184, 0.040)
Live w father	-0.112	(-0.430, 0.206)
Health	-0.047*	(-0.071, -0.023)
Mom edu not know	-0.036	(-0.137, 0.066)
Mom edu less HS	0.007	(-0.075, 0.089)
Mom edu college	0.010	(-0.053, 0.073)
Dad edu not know	-0.023	(-0.130, 0.084)
Dad edu less HS	-0.111*	(-0.200, -0.022)
Dad edu college	0.121*	(0.052, 0.191)
Mom job other	0.047	(-0.039, 0.134)
Mom job professionals	0.162*	(0.085, 0.239)
Mom job office/sales	-0.004	(-0.077, 0.069)
Mom job manual/farmer	0.073	(-0.011, 0.158)
Mom job military	-0.008	(-0.275, 0.259)
Dad job other	0.166	(-0.167, 0.499)
Dad job professionals	0.284	(-0.025, 0.592)
Dad job office/sales	0.413*	(0.095, 0.732)
Dad job manual/farmer	0.194	(-0.103, 0.492)
Dad job military	0.213	(-0.138, 0.564)
Household size	0.012	(-0.007, 0.032)
Contextual effects		
Age	0.003	(-0.077, 0.083)

Continued on next page

Table B.16 – continued

	Main (I)	
Male	0.010	(-0.122, 0.143)
Grade 10	-0.077	(-0.302, 0.148)
Grade 11	0.028	(-0.377, 0.432)
Grade 12	0.076	(-0.570, 0.721)
Hispanic	0.219*	(0.020, 0.418)
Black	-0.020	(-0.167, 0.128)
Asian	-0.100	(-0.506, 0.306)
Other races	0.019	(-0.185, 0.222)
Years in school	-0.020	(-0.108, 0.068)
Live w mother	-0.057	(-0.302, 0.187)
Live w father	0.038	(-0.696, 0.772)
Health	-0.037	(-0.115, 0.041)
Mom edu not know	0.102	(-0.165, 0.369)
Mom edu less HS	-0.178*	(-0.353, -0.003)
Mom edu college	-0.050	(-0.224, 0.125)
Dad edu not know	-0.211*	(-0.399, -0.023)
Dad edu less HS	-0.196*	(-0.387, -0.005)
Dad edu college	0.099	(-0.033, 0.231)
Mom job other	0.072	(-0.121, 0.265)
Mom job professionals	0.087	(-0.055, 0.229)
Mom job office/sales	-0.013	(-0.174, 0.149)
Mom job manual/farmer	0.010	(-0.167, 0.186)
Mom job military	-0.058	(-0.922, 0.807)
Dad job other	-0.011	(-0.728, 0.706)
Dad job professionals	-0.011	(-0.697, 0.674)
Dad job office/sales	-0.112	(-0.800, 0.575)
Dad job manual/farmer	-0.079	(-0.785, 0.627)
Dad job military	0.005	(-0.745, 0.754)
Household size	-0.060*	(-0.114, -0.007)
School characteristics		
Small (1-400)	0.566	(-0.156, 1.287)
Large (1001-4000)	0.080	(-0.276, 0.437)
Urban	-0.131	(-0.511, 0.249)
Rural	0.130	(-0.170, 0.430)
West	-0.398	(-1.110, 0.314)
Midwest	-0.114	(-0.703, 0.474)
South	-0.075	(-0.630, 0.480)

Continued on next page

Table B.16 – continued

	Main (I)	
Private	0.500	(-0.271, 1.270)
Constant	2.403*	(1.213, 3.594)

(I): MLE for strategic binary choice with incomplete information. 95% confidence interval in parentheses, computed from 500 bootstraps for (I). * denotes CI does not include 0.

Table B.17: Science GPAs with both endogenous and contextual peer effects - Point estimates

	Bias correction						No bias correction					
	(I).a		(I).b		(I).c		(II).a		(II).b		(III).c	
Endogenous peer effects (λ)	0.102*	(0.050, 0.155)	0.463*	(0.340, 0.587)	0.486*	(0.329, 0.644)	0.107*	(0.057, 0.156)	0.461*	(0.338, 0.583)	0.486*	(0.354, 0.618)
Own effects												
Age	-0.158*	(-0.198, -0.117)	-0.134*	(-0.168, -0.100)	-0.128*	(-0.158, -0.097)	-0.166*	(-0.202, -0.131)	-0.132*	(-0.164, -0.100)	-0.128*	(-0.158, -0.098)
Male	-0.157*	(-0.199, -0.115)	-0.166*	(-0.207, -0.124)	-0.168*	(-0.210, -0.126)	-0.156*	(-0.198, -0.114)	-0.166*	(-0.207, -0.125)	-0.168*	(-0.210, -0.126)
Grade 10	0.104	(-0.005, 0.212)	0.113*	(0.016, 0.211)	0.113*	(0.022, 0.204)	0.127*	(0.030, 0.225)	0.102*	(0.009, 0.195)	0.114*	(0.020, 0.207)
Grade 11	0.344*	(0.204, 0.484)	0.286*	(0.170, 0.403)	0.291*	(0.183, 0.399)	0.329*	(0.219, 0.439)	0.285*	(0.185, 0.386)	0.291*	(0.189, 0.394)
Grade 12	0.622*	(0.365, 0.879)	0.464*	(0.267, 0.662)	0.484*	(0.332, 0.637)	0.555*	(0.413, 0.697)	0.481*	(0.353, 0.609)	0.484*	(0.360, 0.608)
Hispanic	-0.088*	(-0.153, -0.023)	-0.093*	(-0.149, -0.038)	-0.087*	(-0.141, -0.034)	-0.095*	(-0.156, -0.034)	-0.091*	(-0.145, -0.037)	-0.087*	(-0.141, -0.034)
Black	-0.186*	(-0.279, -0.093)	-0.168*	(-0.251, -0.085)	-0.181*	(-0.261, -0.101)	-0.181*	(-0.268, -0.093)	-0.172*	(-0.253, -0.091)	-0.181*	(-0.259, -0.102)
Asian	0.164*	(0.084, 0.244)	0.208*	(0.140, 0.275)	0.192*	(0.129, 0.254)	0.181*	(0.116, 0.247)	0.201*	(0.142, 0.260)	0.192*	(0.131, 0.253)
Other races	-0.012	(-0.066, 0.042)	-0.020	(-0.069, 0.029)	-0.015	(-0.060, 0.030)	-0.023	(-0.071, 0.024)	-0.018	(-0.064, 0.028)	-0.015	(-0.060, 0.030)
Years in school	0.022*	(0.002, 0.042)	0.019*	(0.001, 0.037)	0.021*	(0.003, 0.038)	0.022*	(0.003, 0.041)	0.019*	(0.002, 0.037)	0.021*	(0.004, 0.038)
Live w mother	0.167*	(0.101, 0.232)	0.162*	(0.101, 0.223)	0.165*	(0.107, 0.223)	0.161*	(0.102, 0.221)	0.164*	(0.106, 0.222)	0.165*	(0.106, 0.223)
Live w father	0.009	(-0.242, 0.261)	-0.003	(-0.256, 0.250)	-0.006	(-0.260, 0.247)	0.002	(-0.245, 0.250)	-0.000	(-0.253, 0.253)	-0.006	(-0.260, 0.247)
Health	-0.111*	(-0.130, -0.091)	-0.108*	(-0.125, -0.091)	-0.105*	(-0.121, -0.088)	-0.115*	(-0.130, -0.099)	-0.107*	(-0.123, -0.091)	-0.105*	(-0.121, -0.089)
Mom edu not know	0.027	(-0.037, 0.091)	0.023	(-0.036, 0.081)	0.020	(-0.036, 0.076)	0.021	(-0.038, 0.081)	0.023	(-0.033, 0.080)	0.020	(-0.036, 0.076)
Mom edu less HS	-0.111*	(-0.156, -0.067)	-0.119*	(-0.163, -0.074)	-0.118*	(-0.163, -0.073)	-0.111*	(-0.154, -0.068)	-0.119*	(-0.162, -0.075)	-0.118*	(-0.162, -0.074)
Mom edu college	0.086*	(0.049, 0.122)	0.086*	(0.053, 0.119)	0.084*	(0.053, 0.115)	0.089*	(0.056, 0.123)	0.086*	(0.056, 0.116)	0.084*	(0.053, 0.115)
Dad edu not know	-0.091*	(-0.156, -0.025)	-0.100*	(-0.162, -0.038)	-0.096*	(-0.160, -0.032)	-0.096*	(-0.161, -0.032)	-0.099*	(-0.161, -0.037)	-0.096*	(-0.159, -0.033)
Dad edu less HS	-0.042	(-0.110, 0.026)	-0.054	(-0.115, 0.007)	-0.051	(-0.109, 0.007)	-0.052	(-0.112, 0.009)	-0.050	(-0.107, 0.007)	-0.051	(-0.108, 0.005)
Dad edu college	0.087*	(0.046, 0.128)	0.095*	(0.055, 0.135)	0.092*	(0.054, 0.130)	0.094*	(0.057, 0.132)	0.092*	(0.054, 0.130)	0.092*	(0.055, 0.129)
Mom job other	-0.103*	(-0.180, -0.027)	-0.084*	(-0.154, -0.014)	-0.085*	(-0.149, -0.022)	-0.091*	(-0.158, -0.023)	-0.085*	(-0.151, -0.019)	-0.085*	(-0.150, -0.021)
Mom job professionals	-0.001	(-0.046, 0.045)	0.013	(-0.026, 0.052)	0.007	(-0.028, 0.041)	0.011	(-0.026, 0.048)	0.009	(-0.026, 0.044)	0.007	(-0.027, 0.041)
Mom job office/sales	-0.049*	(-0.091, -0.007)	-0.039*	(-0.077, -0.001)	-0.041*	(-0.076, -0.006)	-0.042*	(-0.081, -0.004)	-0.040*	(-0.076, -0.003)	-0.041*	(-0.076, -0.005)
Mom job manual/farmer	-0.069*	(-0.130, -0.008)	-0.063*	(-0.119, -0.007)	-0.065*	(-0.119, -0.011)	-0.065*	(-0.125, -0.005)	-0.065*	(-0.120, -0.009)	-0.065*	(-0.120, -0.010)
Mom job military	-0.067	(-0.283, 0.149)	-0.085	(-0.293, 0.123)	-0.093	(-0.306, 0.120)	-0.071	(-0.285, 0.143)	-0.085	(-0.293, 0.123)	-0.093	(-0.306, 0.120)
Dad job other	-0.023	(-0.259, 0.213)	-0.010	(-0.251, 0.230)	-0.010	(-0.251, 0.232)	-0.016	(-0.248, 0.217)	-0.015	(-0.257, 0.227)	-0.009	(-0.251, 0.232)
Dad job professionals	0.094	(-0.154, 0.342)	0.123	(-0.124, 0.371)	0.120	(-0.130, 0.369)	0.118	(-0.126, 0.363)	0.116	(-0.135, 0.367)	0.120	(-0.131, 0.371)
Dad job office/sales	0.055	(-0.190, 0.300)	0.087	(-0.155, 0.328)	0.077	(-0.165, 0.320)	0.079	(-0.160, 0.318)	0.077	(-0.167, 0.321)	0.078	(-0.166, 0.321)
Dad job manual/farmer	0.041	(-0.199, 0.281)	0.065	(-0.178, 0.307)	0.063	(-0.181, 0.308)	0.057	(-0.181, 0.295)	0.060	(-0.185, 0.305)	0.063	(-0.182, 0.309)
Dad job military	0.066	(-0.207, 0.340)	0.068	(-0.204, 0.339)	0.073	(-0.198, 0.344)	0.079	(-0.192, 0.349)	0.062	(-0.211, 0.335)	0.073	(-0.199, 0.345)
Household size	-0.007	(-0.017, 0.004)	-0.005	(-0.014, 0.004)	-0.005	(-0.014, 0.004)	-0.006	(-0.016, 0.004)	-0.005	(-0.014, 0.004)	-0.005	(-0.014, 0.005)
Contextual effects												
Age	-0.028*	(-0.044, -0.013)	-0.089*	(-0.113, -0.066)	-0.092*	(-0.121, -0.063)	-0.028*	(-0.043, -0.012)	-0.090*	(-0.113, -0.067)	-0.092*	(-0.116, -0.068)
Male	0.095*	(0.015, 0.175)	0.180*	(0.108, 0.253)	0.185*	(0.110, 0.260)	0.107*	(0.035, 0.178)	0.180*	(0.113, 0.248)	0.185*	(0.112, 0.257)
Grade 10	-0.010	(-0.103, 0.082)	0.060	(-0.025, 0.146)	0.046	(-0.043, 0.134)	-0.015	(-0.104, 0.074)	0.062	(-0.021, 0.146)	0.046	(-0.043, 0.134)
Grade 11	-0.067	(-0.201, 0.066)	0.066	(-0.050, 0.183)	0.059	(-0.055, 0.172)	-0.053	(-0.161, 0.054)	0.065	(-0.037, 0.167)	0.059	(-0.048, 0.166)
Grade 12	-0.050	(-0.211, 0.111)	0.118	(-0.012, 0.249)	0.107	(-0.022, 0.235)	-0.028	(-0.152, 0.095)	0.116*	(0.008, 0.224)	0.107	(-0.008, 0.222)

Continued on next page

Table B.17 – continued

	Bias correction						No bias correction					
	(I).a		(I).b		(I).c		(II).a		(II).b		(III).c	
Hispanic	-0.066	(-0.262, 0.130)	0.002	(-0.150, 0.154)	-0.013	(-0.171, 0.144)	-0.056	(-0.248, 0.137)	-0.004	(-0.156, 0.148)	-0.013	(-0.171, 0.145)
Black	-0.045	(-0.159, 0.070)	0.026	(-0.061, 0.113)	0.049	(-0.043, 0.142)	-0.040	(-0.153, 0.072)	0.026	(-0.060, 0.113)	0.049	(-0.040, 0.139)
Asian	0.037	(-0.116, 0.189)	-0.097	(-0.239, 0.044)	-0.082	(-0.212, 0.048)	0.027	(-0.116, 0.170)	-0.096	(-0.232, 0.040)	-0.082	(-0.210, 0.046)
Other races	-0.073	(-0.172, 0.026)	-0.057	(-0.149, 0.036)	-0.061	(-0.153, 0.031)	-0.078	(-0.173, 0.018)	-0.057	(-0.149, 0.034)	-0.061	(-0.154, 0.032)
Years in school	0.029	(-0.017, 0.075)	0.014	(-0.020, 0.048)	0.010	(-0.024, 0.045)	0.027	(-0.016, 0.070)	0.015	(-0.018, 0.047)	0.010	(-0.024, 0.044)
Live w mother	0.168*	(0.049, 0.287)	0.034	(-0.092, 0.160)	0.026	(-0.111, 0.163)	0.170*	(0.054, 0.286)	0.036	(-0.089, 0.161)	0.026	(-0.109, 0.161)
Live w father	0.016	(-0.475, 0.507)	0.141	(-0.349, 0.631)	0.100	(-0.396, 0.597)	0.003	(-0.472, 0.479)	0.136	(-0.349, 0.620)	0.100	(-0.392, 0.593)
Health	-0.081*	(-0.115, -0.047)	-0.036*	(-0.069, -0.003)	-0.041*	(-0.076, -0.005)	-0.084*	(-0.115, -0.052)	-0.036*	(-0.067, -0.004)	-0.041*	(-0.076, -0.006)
Mom edu not know	-0.029	(-0.177, 0.120)	-0.071	(-0.188, 0.047)	-0.016	(-0.147, 0.116)	-0.046	(-0.186, 0.094)	-0.073	(-0.189, 0.043)	-0.016	(-0.146, 0.114)
Mom edu less HS	0.048	(-0.074, 0.170)	0.100	(-0.027, 0.227)	0.123	(-0.007, 0.253)	0.039	(-0.077, 0.154)	0.105	(-0.020, 0.230)	0.123	(-0.005, 0.250)
Mom edu college	0.102*	(0.030, 0.173)	0.040	(-0.026, 0.106)	0.038	(-0.031, 0.107)	0.093*	(0.028, 0.159)	0.042	(-0.022, 0.106)	0.038	(-0.029, 0.105)
Dad edu not know	0.071	(-0.051, 0.192)	0.107	(-0.002, 0.216)	0.089	(-0.031, 0.210)	0.064	(-0.051, 0.179)	0.111*	(0.005, 0.217)	0.089	(-0.031, 0.210)
Dad edu less HS	0.047	(-0.094, 0.189)	0.007	(-0.111, 0.126)	0.037	(-0.088, 0.161)	0.033	(-0.098, 0.164)	0.014	(-0.103, 0.130)	0.037	(-0.090, 0.163)
Dad edu college	0.048	(-0.050, 0.146)	0.018	(-0.072, 0.108)	0.007	(-0.083, 0.098)	0.066	(-0.024, 0.156)	0.016	(-0.072, 0.103)	0.007	(-0.081, 0.096)
Mom job other	-0.070	(-0.195, 0.055)	-0.025	(-0.129, 0.078)	-0.027	(-0.143, 0.088)	-0.076	(-0.200, 0.047)	-0.028	(-0.131, 0.074)	-0.027	(-0.140, 0.085)
Mom job professionals	0.018	(-0.070, 0.105)	-0.010	(-0.089, 0.068)	0.006	(-0.074, 0.086)	0.024	(-0.061, 0.109)	-0.014	(-0.091, 0.063)	0.006	(-0.074, 0.086)
Mom job office/sales	-0.072	(-0.165, 0.021)	-0.082	(-0.165, 0.001)	-0.065	(-0.144, 0.014)	-0.072	(-0.162, 0.019)	-0.084*	(-0.164, -0.003)	-0.065	(-0.144, 0.014)
Mom job manual/farmer	-0.085	(-0.217, 0.047)	-0.032	(-0.147, 0.084)	-0.024	(-0.147, 0.100)	-0.094	(-0.224, 0.036)	-0.033	(-0.145, 0.079)	-0.024	(-0.145, 0.097)
Mom job military	0.226	(-0.087, 0.538)	0.126	(-0.235, 0.488)	0.145	(-0.185, 0.476)	0.214	(-0.081, 0.509)	0.126	(-0.231, 0.483)	0.146	(-0.180, 0.471)
Dad job other	0.037	(-0.455, 0.529)	-0.087	(-0.579, 0.406)	-0.055	(-0.550, 0.440)	0.039	(-0.444, 0.521)	-0.082	(-0.573, 0.408)	-0.055	(-0.549, 0.439)
Dad job professionals	0.112	(-0.378, 0.602)	-0.104	(-0.594, 0.386)	-0.054	(-0.552, 0.443)	0.123	(-0.355, 0.601)	-0.098	(-0.585, 0.389)	-0.054	(-0.547, 0.438)
Dad job office/sales	0.092	(-0.380, 0.565)	-0.101	(-0.565, 0.363)	-0.035	(-0.511, 0.441)	0.107	(-0.352, 0.565)	-0.094	(-0.554, 0.366)	-0.035	(-0.506, 0.437)
Dad job manual/farmer	-0.005	(-0.485, 0.474)	-0.180	(-0.655, 0.295)	-0.124	(-0.605, 0.358)	0.005	(-0.462, 0.473)	-0.174	(-0.645, 0.298)	-0.124	(-0.602, 0.354)
Dad job military	-0.007	(-0.534, 0.520)	-0.118	(-0.632, 0.395)	-0.123	(-0.654, 0.408)	-0.004	(-0.517, 0.510)	-0.114	(-0.624, 0.396)	-0.123	(-0.651, 0.405)
Household size	0.016	(-0.013, 0.045)	0.007	(-0.023, 0.037)	0.002	(-0.026, 0.031)	0.014	(-0.013, 0.041)	0.009	(-0.020, 0.037)	0.002	(-0.027, 0.031)
School characteristics												
Small (1-400)	-0.149	(-0.556, 0.258)	-0.078	(-0.352, 0.196)	-0.084	(-0.326, 0.159)	-0.107	(-0.493, 0.278)	-0.090	(-0.348, 0.167)	-0.084	(-0.327, 0.160)
Large (1001-4000)	-0.070	(-0.225, 0.085)	-0.029	(-0.127, 0.069)	-0.029	(-0.118, 0.060)	-0.059	(-0.207, 0.089)	-0.031	(-0.126, 0.064)	-0.029	(-0.119, 0.061)
Urban	0.030	(-0.182, 0.243)	0.025	(-0.113, 0.163)	0.025	(-0.101, 0.152)	0.042	(-0.154, 0.238)	0.029	(-0.101, 0.159)	0.025	(-0.100, 0.151)
Rural	0.000	(-0.280, 0.281)	0.017	(-0.160, 0.194)	0.012	(-0.146, 0.170)	0.021	(-0.239, 0.282)	0.014	(-0.150, 0.177)	0.012	(-0.143, 0.167)
West	0.224	(-0.075, 0.523)	0.093	(-0.118, 0.305)	0.100	(-0.057, 0.257)	0.178	(-0.047, 0.402)	0.106	(-0.051, 0.263)	0.100	(-0.052, 0.252)
Midwest	0.159	(-0.057, 0.375)	0.080	(-0.062, 0.221)	0.079	(-0.033, 0.191)	0.127	(-0.055, 0.308)	0.083	(-0.034, 0.200)	0.079	(-0.033, 0.191)
South	0.083	(-0.129, 0.294)	0.043	(-0.091, 0.178)	0.045	(-0.068, 0.158)	0.067	(-0.120, 0.254)	0.046	(-0.071, 0.163)	0.045	(-0.068, 0.159)
Private	0.132	(-0.229, 0.493)	0.062	(-0.168, 0.293)	0.045	(-0.163, 0.252)	0.121	(-0.225, 0.468)	0.054	(-0.165, 0.272)	0.045	(-0.161, 0.251)
Constant	5.111*	(4.543, 5.679)	4.679*	(4.210, 5.148)	4.602*	(4.156, 5.049)	5.175*	(4.644, 5.707)	4.672*	(4.219, 5.125)	4.603*	(4.163, 5.042)
Selection	-0.337	(-1.226, 0.553)	0.100	(-0.532, 0.733)	-0.001	(-0.303, 0.301)						

a: Dynamic peer structure for endogenous peer effects, fixed peer structure for contextual peer effects. b: Dynamic peer structure for both types of peer effects. c: Fixed peer structure for both types of peer effects. All are estimated by NLS. 95% confidence interval in parentheses, computed from 500 bootstraps. * denotes CI does not include 0.

Table B.18: Science GPAs with both endogenous and contextual peer effects - Marginal effects

	(I).a				(I).b				(I).c			
	Average own effect		Average total effect		Average own effect		Average total effect		Average own effect		Average total effect	
Age	-0.158*	(-0.238, -0.157)	-0.197*	(-0.252, -0.143)	-0.143*	(-0.394, -0.326)	-0.360*	(-0.466, -0.253)	-0.143*	(-0.394, -0.326)	-0.360*	(-0.466, -0.253)
Male	-0.156*	(-0.131, -0.047)	-0.089*	(-0.163, -0.015)	-0.160	(-0.053, 0.030)	-0.011	(-0.110, 0.088)	-0.160	(-0.053, 0.030)	-0.011	(-0.110, 0.088)
Grade 10	0.104	(-0.004, 0.213)	0.104	(-0.005, 0.214)	0.120*	(0.184, 0.378)	0.281*	(0.137, 0.425)	0.120*	(0.184, 0.378)	0.281*	(0.137, 0.425)
Grade 11	0.344*	(0.179, 0.458)	0.318*	(0.181, 0.456)	0.298*	(0.467, 0.699)	0.583*	(0.353, 0.812)	0.298*	(0.467, 0.699)	0.583*	(0.353, 0.812)
Grade 12	0.622*	(0.381, 0.895)	0.638*	(0.386, 0.890)	0.484*	(0.762, 1.161)	0.961*	(0.583, 1.340)	0.484*	(0.762, 1.161)	0.961*	(0.583, 1.340)
Hispanic	-0.089*	(-0.220, -0.087)	-0.154	(-0.375, 0.067)	-0.096*	(-0.217, -0.093)	-0.155	(-0.442, 0.132)	-0.096*	(-0.217, -0.093)	-0.155	(-0.442, 0.132)
Black	-0.187*	(-0.336, -0.150)	-0.243*	(-0.389, -0.097)	-0.171*	(-0.329, -0.161)	-0.245*	(-0.398, -0.092)	-0.171*	(-0.329, -0.161)	-0.245*	(-0.398, -0.092)
Asian	0.164*	(0.131, 0.292)	0.211*	(0.039, 0.383)	0.208*	(0.138, 0.276)	0.207	(-0.027, 0.441)	0.208*	(0.138, 0.276)	0.207	(-0.027, 0.441)
Other races	-0.013*	(-0.131, -0.022)	-0.077	(-0.198, 0.045)	-0.024*	(-0.170, -0.067)	-0.118	(-0.288, 0.051)	-0.024*	(-0.170, -0.067)	-0.118	(-0.288, 0.051)
Years in school	0.022*	(0.029, 0.070)	0.049	(-0.006, 0.105)	0.020*	(0.034, 0.072)	0.053	(-0.012, 0.118)	0.020*	(0.034, 0.072)	0.053	(-0.012, 0.118)
Live w mother	0.169*	(0.264, 0.395)	0.329*	(0.199, 0.460)	0.169*	(0.263, 0.387)	0.325*	(0.139, 0.511)	0.169*	(0.263, 0.387)	0.325*	(0.139, 0.511)
Live w father	0.009	(-0.228, 0.276)	0.024	(-0.476, 0.524)	0.005	(-0.048, 0.459)	0.205	(-0.538, 0.948)	0.005	(-0.048, 0.459)	0.205	(-0.538, 0.948)
Health	-0.111*	(-0.211, -0.172)	-0.191*	(-0.235, -0.147)	-0.113*	(-0.254, -0.219)	-0.236*	(-0.292, -0.180)	-0.113*	(-0.254, -0.219)	-0.236*	(-0.292, -0.180)
Mom edu not know	0.027	(-0.060, 0.069)	0.005	(-0.174, 0.184)	0.019*	(-0.130, -0.005)	-0.067	(-0.302, 0.167)	0.019*	(-0.130, -0.005)	-0.067	(-0.302, 0.167)
Mom edu less HS	-0.111*	(-0.125, -0.036)	-0.081	(-0.201, 0.040)	-0.116*	(-0.097, -0.007)	-0.052	(-0.247, 0.143)	-0.116*	(-0.097, -0.007)	-0.052	(-0.247, 0.143)
Mom edu college	0.087*	(0.146, 0.219)	0.182*	(0.105, 0.260)	0.091*	(0.172, 0.238)	0.205*	(0.088, 0.322)	0.091*	(0.172, 0.238)	0.205*	(0.088, 0.322)
Dad edu not know	-0.090	(-0.104, 0.028)	-0.038	(-0.171, 0.095)	-0.096	(-0.074, 0.054)	-0.010	(-0.198, 0.178)	-0.096	(-0.074, 0.054)	-0.010	(-0.198, 0.178)
Dad edu less HS	-0.041	(-0.073, 0.064)	-0.005	(-0.170, 0.161)	-0.055*	(-0.143, -0.016)	-0.080	(-0.292, 0.132)	-0.055*	(-0.143, -0.016)	-0.080	(-0.292, 0.132)
Dad edu college	0.088*	(0.096, 0.179)	0.137*	(0.031, 0.244)	0.098*	(0.146, 0.228)	0.187*	(0.046, 0.328)	0.098*	(0.146, 0.228)	0.187*	(0.046, 0.328)
Mom job other	-0.104*	(-0.251, -0.097)	-0.174*	(-0.329, -0.019)	-0.087*	(-0.252, -0.107)	-0.179	(-0.378, 0.019)	-0.087*	(-0.252, -0.107)	-0.179	(-0.378, 0.019)
Mom job professionals	-0.000	(-0.032, 0.061)	0.015	(-0.098, 0.127)	0.013	(-0.035, 0.049)	0.007	(-0.145, 0.159)	0.013	(-0.035, 0.049)	0.007	(-0.145, 0.159)
Mom job office/sales	-0.050*	(-0.159, -0.073)	-0.116*	(-0.227, -0.005)	-0.045*	(-0.229, -0.149)	-0.189*	(-0.343, -0.035)	-0.045*	(-0.229, -0.149)	-0.189*	(-0.343, -0.035)
Mom job manual/farmer	-0.070*	(-0.212, -0.088)	-0.150	(-0.307, 0.008)	-0.066*	(-0.212, -0.095)	-0.153	(-0.358, 0.052)	-0.066*	(-0.212, -0.095)	-0.153	(-0.358, 0.052)
Mom job military	-0.065	(-0.094, 0.339)	0.123	(-0.248, 0.494)	-0.080	(-0.165, 0.254)	0.045	(-0.529, 0.619)	-0.080	(-0.165, 0.254)	0.045	(-0.529, 0.619)
Dad job other	-0.023	(-0.228, 0.243)	0.007	(-0.479, 0.493)	-0.016	(-0.387, 0.094)	-0.147	(-0.883, 0.590)	-0.016	(-0.387, 0.094)	-0.147	(-0.883, 0.590)
Dad job professionals	0.095	(-0.048, 0.449)	0.201	(-0.318, 0.720)	0.120	(-0.196, 0.303)	0.054	(-0.708, 0.815)	0.120	(-0.196, 0.303)	0.054	(-0.708, 0.815)
Dad job office/sales	0.056	(-0.105, 0.387)	0.141	(-0.376, 0.658)	0.083	(-0.248, 0.241)	-0.004	(-0.743, 0.736)	0.083	(-0.248, 0.241)	-0.004	(-0.743, 0.736)
Dad job manual/farmer	0.041	(-0.200, 0.281)	0.040	(-0.445, 0.526)	0.056	(-0.402, 0.084)	-0.159	(-0.876, 0.558)	0.056	(-0.402, 0.084)	-0.159	(-0.876, 0.558)
Dad job military	0.066	(-0.208, 0.341)	0.066	(-0.515, 0.648)	0.063	(-0.338, 0.215)	-0.062	(-0.905, 0.782)	0.063	(-0.338, 0.215)	-0.062	(-0.905, 0.782)
Household size	-0.007	(-0.004, 0.017)	0.006	(-0.023, 0.036)	-0.004	(-0.007, 0.012)	0.003	(-0.045, 0.050)	-0.004	(-0.007, 0.012)	0.003	(-0.045, 0.050)

Marginal effects for (I) in Table B.17. Marginal effects for students selecting the subject only. Marginal effects average within in each subnetwork then average over subnetworks. a: Dynamic peer structure for endogenous peer effects, fixed peer structure for contextual peer effects. b: Dynamic peer structure for both types of peer effects. c: Fixed peer structure for both types of peer effects. 95% confidence interval in parentheses, computed from 500 bootstraps. * denotes CI does not include 0.

Table B.19: Social Studies sample: Selection to take a course in Social Studies with contextual peer effects

	Main (I)	
Endogenous peer effects (δ)	2.223*	(0.440, 4.006)
Own effects		
Age	0.031	(-0.052, 0.113)
Male	0.057*	(0.016, 0.097)
Grade 10	0.113	(-0.306, 0.532)
Grade 11	0.806*	(0.361, 1.251)
Grade 12	0.183	(-0.287, 0.654)
Hispanic	0.039	(-0.043, 0.121)
Black	0.093	(-0.015, 0.201)
Asian	0.018	(-0.122, 0.158)
Other races	-0.110*	(-0.187, -0.033)
Years in school	-0.017	(-0.057, 0.022)
Live w mother	0.015	(-0.075, 0.105)
Live w father	-0.110	(-0.422, 0.201)
Health	-0.011	(-0.033, 0.011)
Mom edu not know	-0.000	(-0.089, 0.089)
Mom edu less HS	0.134*	(0.053, 0.215)
Mom edu college	0.019	(-0.046, 0.084)
Dad edu not know	-0.049	(-0.159, 0.061)
Dad edu less HS	-0.004	(-0.089, 0.082)
Dad edu college	0.049	(-0.003, 0.102)
Mom job other	-0.046	(-0.123, 0.031)
Mom job professionals	0.041	(-0.025, 0.107)
Mom job office/sales	0.044	(-0.011, 0.100)
Mom job manual/farmer	0.022	(-0.050, 0.094)
Mom job military	0.026	(-0.288, 0.340)
Dad job other	0.133	(-0.202, 0.469)
Dad job professionals	0.165	(-0.169, 0.498)
Dad job office/sales	0.175	(-0.166, 0.516)
Dad job manual/farmer	0.140	(-0.173, 0.453)
Dad job military	0.120	(-0.251, 0.492)
Household size	0.011	(-0.012, 0.033)
Contextual effects		
Age	-0.093	(-0.188, 0.002)
Male	-0.004	(-0.115, 0.108)
Grade 10	-0.020	(-0.325, 0.286)
Grade 11	-0.369*	(-0.729, -0.010)

Continued on next page

Table B.19 – continued

	Main (I)	
Grade 12	0.130	(-0.344, 0.604)
Hispanic	-0.040	(-0.218, 0.137)
Black	0.008	(-0.125, 0.140)
Asian	-0.174	(-0.441, 0.093)
Other races	0.012	(-0.243, 0.268)
Years in school	-0.022	(-0.116, 0.073)
Live w mother	-0.090	(-0.346, 0.165)
Live w father	-0.558	(-1.369, 0.253)
Health	0.041	(-0.021, 0.103)
Mom edu not know	-0.051	(-0.305, 0.203)
Mom edu less HS	-0.086	(-0.263, 0.090)
Mom edu college	0.032	(-0.152, 0.216)
Dad edu not know	0.137	(-0.185, 0.459)
Dad edu less HS	-0.101	(-0.353, 0.151)
Dad edu college	0.024	(-0.139, 0.186)
Mom job other	0.006	(-0.184, 0.196)
Mom job professionals	0.006	(-0.180, 0.191)
Mom job office/sales	-0.032	(-0.224, 0.161)
Mom job manual/farmer	0.159	(-0.081, 0.399)
Mom job military	0.047	(-0.666, 0.760)
Dad job other	0.383	(-0.382, 1.149)
Dad job professionals	0.605	(-0.147, 1.356)
Dad job office/sales	0.556	(-0.172, 1.284)
Dad job manual/farmer	0.455	(-0.271, 1.181)
Dad job military	0.698	(-0.048, 1.443)
Household size	-0.030	(-0.074, 0.013)
School characteristics		
Small (1-400)	-0.042	(-0.474, 0.390)
Large (1001-4000)	-0.170	(-0.400, 0.060)
Urban	0.020	(-0.178, 0.218)
Rural	-0.080	(-0.456, 0.296)
West	-0.245	(-2.375, 1.886)
Midwest	-0.393	(-2.489, 1.702)
South	-0.244	(-2.344, 1.855)
Private	-0.120	(-1.889, 1.650)
Constant	0.424	(-1.960, 2.809)

(I): MLE for strategic binary choice with incomplete information. 95% confidence interval in parentheses, computed from 250 bootstraps. * denotes CI does not include 0.

Table B.20: Social Studies GPAs with both endogenous and contextual peer effects - Point estimates

	Bias correction						No bias correction					
	(I).a		(I).b		(I).c		(II).a		(II).b		(III).c	
Endogenous peer effects (λ)	0.036	(-0.020, 0.092)	0.392*	(0.258, 0.527)	0.436*	(0.296, 0.576)	0.045	(-0.006, 0.095)	0.399*	(0.272, 0.526)	0.434*	(0.318, 0.550)
Own effects												
Age	-0.185*	(-0.231, -0.140)	-0.146*	(-0.186, -0.106)	-0.138*	(-0.179, -0.098)	-0.179*	(-0.222, -0.135)	-0.145*	(-0.186, -0.105)	-0.140*	(-0.179, -0.101)
Male	-0.110*	(-0.148, -0.073)	-0.114*	(-0.151, -0.076)	-0.117*	(-0.154, -0.081)	-0.107*	(-0.144, -0.070)	-0.112*	(-0.149, -0.076)	-0.115*	(-0.152, -0.077)
Grade 10	0.124	(-0.016, 0.264)	0.107	(-0.018, 0.232)	0.099	(-0.026, 0.224)	0.118*	(0.006, 0.231)	0.111	(-0.008, 0.230)	0.106	(-0.012, 0.224)
Grade 11	0.302*	(0.132, 0.471)	0.271*	(0.126, 0.416)	0.257*	(0.117, 0.397)	0.330*	(0.204, 0.456)	0.292*	(0.168, 0.415)	0.288*	(0.155, 0.422)
Grade 12	0.678*	(0.492, 0.864)	0.564*	(0.396, 0.732)	0.561*	(0.393, 0.729)	0.647*	(0.487, 0.807)	0.571*	(0.406, 0.735)	0.576*	(0.415, 0.736)
Hispanic	-0.102*	(-0.166, -0.038)	-0.100*	(-0.158, -0.041)	-0.086*	(-0.142, -0.030)	-0.098*	(-0.157, -0.038)	-0.099*	(-0.157, -0.040)	-0.085*	(-0.139, -0.030)
Black	-0.167*	(-0.247, -0.087)	-0.170*	(-0.259, -0.082)	-0.154*	(-0.233, -0.074)	-0.163*	(-0.244, -0.082)	-0.167*	(-0.255, -0.079)	-0.151*	(-0.232, -0.069)
Asian	0.179*	(0.110, 0.247)	0.188*	(0.122, 0.254)	0.192*	(0.128, 0.255)	0.180*	(0.111, 0.248)	0.188*	(0.122, 0.253)	0.191*	(0.127, 0.255)
Other races	-0.028	(-0.084, 0.028)	-0.031	(-0.084, 0.021)	-0.027	(-0.079, 0.024)	-0.039	(-0.094, 0.015)	-0.035	(-0.087, 0.016)	-0.031	(-0.085, 0.022)
Years in school	0.014	(-0.009, 0.037)	0.016	(-0.005, 0.037)	0.015	(-0.004, 0.034)	0.016	(-0.006, 0.037)	0.015	(-0.005, 0.036)	0.014	(-0.005, 0.033)
Live w mother	0.152*	(0.087, 0.217)	0.156*	(0.097, 0.216)	0.154*	(0.095, 0.213)	0.151*	(0.086, 0.215)	0.156*	(0.098, 0.214)	0.153*	(0.094, 0.211)
Live w father	0.034	(-0.213, 0.280)	-0.010	(-0.249, 0.230)	-0.000	(-0.246, 0.246)	0.005	(-0.239, 0.248)	-0.016	(-0.256, 0.223)	-0.009	(-0.256, 0.239)
Health	-0.118*	(-0.132, -0.103)	-0.111*	(-0.125, -0.098)	-0.108*	(-0.121, -0.095)	-0.119*	(-0.132, -0.106)	-0.112*	(-0.125, -0.099)	-0.108*	(-0.121, -0.096)
Mom edu not know	-0.031	(-0.080, 0.017)	-0.030	(-0.078, 0.019)	-0.033	(-0.079, 0.013)	-0.028	(-0.076, 0.020)	-0.030	(-0.077, 0.018)	-0.033	(-0.082, 0.015)
Mom edu less HS	-0.119*	(-0.173, -0.066)	-0.125*	(-0.174, -0.076)	-0.121*	(-0.172, -0.071)	-0.114*	(-0.164, -0.064)	-0.121*	(-0.170, -0.072)	-0.116*	(-0.166, -0.066)
Mom edu college	0.108*	(0.078, 0.139)	0.108*	(0.078, 0.138)	0.104*	(0.075, 0.134)	0.111*	(0.082, 0.140)	0.109*	(0.079, 0.138)	0.105*	(0.076, 0.134)
Dad edu not know	-0.132*	(-0.203, -0.061)	-0.139*	(-0.203, -0.075)	-0.137*	(-0.198, -0.076)	-0.140*	(-0.202, -0.078)	-0.140*	(-0.200, -0.081)	-0.139*	(-0.198, -0.079)
Dad edu less HS	-0.061	(-0.123, 0.002)	-0.056	(-0.112, 0.001)	-0.055*	(-0.110, -0.000)	-0.062*	(-0.122, -0.001)	-0.056*	(-0.110, -0.001)	-0.055	(-0.111, 0.001)
Dad edu college	0.143*	(0.102, 0.184)	0.141*	(0.104, 0.179)	0.136*	(0.099, 0.174)	0.146*	(0.109, 0.184)	0.143*	(0.105, 0.180)	0.138*	(0.101, 0.174)
Mom job other	-0.141*	(-0.211, -0.070)	-0.143*	(-0.213, -0.073)	-0.138*	(-0.205, -0.070)	-0.145*	(-0.218, -0.072)	-0.144*	(-0.213, -0.075)	-0.139*	(-0.209, -0.068)
Mom job professionals	0.002	(-0.046, 0.050)	-0.002	(-0.046, 0.042)	-0.001	(-0.044, 0.042)	0.006	(-0.040, 0.051)	-0.001	(-0.044, 0.042)	0.000	(-0.042, 0.042)
Mom job office/sales	-0.046	(-0.097, 0.005)	-0.046	(-0.095, 0.003)	-0.046	(-0.095, 0.002)	-0.043	(-0.094, 0.008)	-0.045	(-0.094, 0.004)	-0.045	(-0.092, 0.002)
Mom job manual/farmer	-0.069*	(-0.127, -0.010)	-0.069*	(-0.120, -0.018)	-0.065*	(-0.116, -0.014)	-0.067*	(-0.124, -0.009)	-0.069*	(-0.119, -0.018)	-0.065*	(-0.118, -0.011)
Mom job military	-0.127	(-0.274, 0.021)	-0.126	(-0.268, 0.016)	-0.132	(-0.275, 0.012)	-0.117	(-0.251, 0.017)	-0.125	(-0.264, 0.014)	-0.128	(-0.261, 0.005)
Dad job other	-0.076	(-0.314, 0.161)	-0.033	(-0.268, 0.202)	-0.046	(-0.286, 0.195)	-0.050	(-0.286, 0.186)	-0.026	(-0.260, 0.209)	-0.038	(-0.280, 0.204)
Dad job professionals	0.059	(-0.199, 0.317)	0.103	(-0.144, 0.350)	0.087	(-0.168, 0.343)	0.093	(-0.163, 0.349)	0.111	(-0.138, 0.360)	0.098	(-0.158, 0.355)
Dad job office/sales	0.078	(-0.187, 0.343)	0.120	(-0.137, 0.377)	0.106	(-0.162, 0.373)	0.116	(-0.144, 0.376)	0.129	(-0.130, 0.387)	0.117	(-0.146, 0.379)
Dad job manual/farmer	-0.017	(-0.264, 0.230)	0.031	(-0.212, 0.273)	0.018	(-0.232, 0.269)	0.016	(-0.228, 0.261)	0.038	(-0.205, 0.282)	0.028	(-0.221, 0.277)
Dad job military	0.065	(-0.198, 0.328)	0.102	(-0.153, 0.358)	0.092	(-0.174, 0.357)	0.102	(-0.153, 0.356)	0.110	(-0.147, 0.366)	0.101	(-0.158, 0.359)
Household size	-0.005	(-0.021, 0.011)	-0.002	(-0.016, 0.012)	-0.002	(-0.016, 0.011)	-0.004	(-0.018, 0.010)	-0.002	(-0.016, 0.012)	-0.002	(-0.015, 0.011)
Contextual effects												
Age	-0.004	(-0.021, 0.013)	-0.066*	(-0.097, -0.035)	-0.074*	(-0.105, -0.042)	-0.005	(-0.021, 0.010)	-0.067*	(-0.097, -0.036)	-0.073*	(-0.100, -0.045)
Male	0.036	(-0.026, 0.097)	0.093*	(0.026, 0.160)	0.102*	(0.027, 0.177)	0.036	(-0.021, 0.094)	0.095*	(0.030, 0.161)	0.102*	(0.032, 0.172)
Grade 10	-0.024	(-0.121, 0.073)	0.040	(-0.060, 0.140)	0.051	(-0.052, 0.155)	-0.013	(-0.108, 0.081)	0.040	(-0.061, 0.140)	0.049	(-0.057, 0.154)
Grade 11	-0.012	(-0.129, 0.106)	0.089	(-0.027, 0.205)	0.106	(-0.025, 0.236)	-0.010	(-0.118, 0.098)	0.089	(-0.024, 0.201)	0.099	(-0.031, 0.229)
Grade 12	-0.062	(-0.200, 0.076)	0.068	(-0.063, 0.200)	0.054	(-0.080, 0.189)	-0.054	(-0.179, 0.072)	0.071	(-0.054, 0.197)	0.059	(-0.075, 0.193)

Continued on next page

Table B.20 – continued

	Bias correction						No bias correction					
	(I).a		(I).b		(I).c		(II).a		(II).b		(III).c	
Hispanic	-0.149	(-0.304, 0.006)	-0.062	(-0.185, 0.061)	-0.079	(-0.225, 0.067)	-0.153*	(-0.304, -0.002)	-0.063	(-0.184, 0.059)	-0.082	(-0.226, 0.061)
Black	-0.108*	(-0.200, -0.016)	-0.014	(-0.089, 0.061)	-0.017	(-0.096, 0.061)	-0.098*	(-0.189, -0.007)	-0.008	(-0.079, 0.064)	-0.010	(-0.090, 0.070)
Asian	0.068	(-0.092, 0.228)	-0.031	(-0.184, 0.121)	-0.052	(-0.206, 0.102)	0.061	(-0.084, 0.207)	-0.041	(-0.189, 0.106)	-0.062	(-0.204, 0.081)
Other races	-0.014	(-0.123, 0.096)	-0.021	(-0.114, 0.071)	-0.030	(-0.135, 0.075)	-0.039	(-0.150, 0.072)	-0.025	(-0.113, 0.064)	-0.033	(-0.144, 0.078)
Years in school	0.002	(-0.052, 0.055)	-0.001	(-0.043, 0.042)	0.001	(-0.037, 0.040)	0.005	(-0.042, 0.053)	-0.002	(-0.041, 0.037)	-0.001	(-0.039, 0.037)
Live w mother	0.193*	(0.065, 0.322)	0.038	(-0.105, 0.181)	-0.004	(-0.136, 0.128)	0.153*	(0.032, 0.274)	0.033	(-0.098, 0.163)	-0.001	(-0.128, 0.126)
Live w father	0.224	(-0.284, 0.733)	0.482	(-0.161, 1.125)	0.283	(-0.278, 0.845)	0.151	(-0.322, 0.624)	0.458	(-0.174, 1.090)	0.248	(-0.292, 0.789)
Health	-0.105*	(-0.146, -0.065)	-0.057*	(-0.092, -0.022)	-0.064*	(-0.101, -0.028)	-0.104*	(-0.143, -0.066)	-0.055*	(-0.090, -0.020)	-0.061*	(-0.097, -0.026)
Mom edu not know	0.016	(-0.091, 0.122)	-0.007	(-0.128, 0.113)	0.062	(-0.038, 0.161)	0.017	(-0.075, 0.109)	-0.010	(-0.125, 0.106)	0.057	(-0.040, 0.154)
Mom edu less HS	0.016	(-0.097, 0.130)	0.094	(-0.012, 0.199)	0.107	(-0.008, 0.221)	0.018	(-0.088, 0.123)	0.094	(-0.010, 0.197)	0.103	(-0.003, 0.209)
Mom edu college	0.083	(-0.009, 0.176)	0.012	(-0.065, 0.090)	0.013	(-0.074, 0.101)	0.076	(-0.017, 0.169)	0.014	(-0.063, 0.091)	0.018	(-0.073, 0.109)
Dad edu not know	-0.022	(-0.165, 0.120)	0.034	(-0.097, 0.165)	0.058	(-0.074, 0.190)	-0.013	(-0.140, 0.114)	0.038	(-0.090, 0.166)	0.064	(-0.062, 0.191)
Dad edu less HS	-0.118	(-0.276, 0.039)	-0.093	(-0.221, 0.034)	-0.064	(-0.194, 0.065)	-0.130	(-0.273, 0.013)	-0.096	(-0.216, 0.024)	-0.069	(-0.202, 0.065)
Dad edu college	0.127*	(0.035, 0.220)	0.060	(-0.023, 0.143)	0.062	(-0.024, 0.147)	0.138*	(0.047, 0.228)	0.061	(-0.020, 0.143)	0.065	(-0.024, 0.153)
Mom job other	-0.128*	(-0.241, -0.015)	-0.037	(-0.154, 0.080)	-0.030	(-0.148, 0.089)	-0.114*	(-0.224, -0.004)	-0.036	(-0.148, 0.075)	-0.031	(-0.141, 0.079)
Mom job professionals	0.044	(-0.064, 0.152)	0.037	(-0.058, 0.132)	0.056	(-0.047, 0.160)	0.080	(-0.030, 0.190)	0.039	(-0.051, 0.129)	0.058	(-0.047, 0.162)
Mom job office/sales	-0.022	(-0.125, 0.081)	-0.031	(-0.124, 0.062)	-0.006	(-0.107, 0.096)	-0.010	(-0.114, 0.094)	-0.031	(-0.123, 0.061)	-0.009	(-0.109, 0.092)
Mom job manual/farmer	-0.101	(-0.244, 0.041)	-0.036	(-0.166, 0.094)	-0.019	(-0.148, 0.111)	-0.065	(-0.196, 0.065)	-0.031	(-0.152, 0.090)	-0.015	(-0.138, 0.107)
Mom job military	-0.133	(-0.557, 0.292)	-0.099	(-0.484, 0.285)	-0.056	(-0.418, 0.306)	-0.062	(-0.491, 0.366)	-0.098	(-0.468, 0.273)	-0.063	(-0.439, 0.313)
Dad job other	-0.209	(-0.734, 0.316)	-0.440	(-1.079, 0.200)	-0.231	(-0.788, 0.325)	-0.124	(-0.602, 0.355)	-0.419	(-1.045, 0.207)	-0.204	(-0.739, 0.330)
Dad job professionals	-0.123	(-0.623, 0.377)	-0.467	(-1.091, 0.157)	-0.251	(-0.789, 0.287)	-0.038	(-0.501, 0.426)	-0.441	(-1.051, 0.170)	-0.212	(-0.735, 0.312)
Dad job office/sales	-0.059	(-0.572, 0.455)	-0.380	(-1.033, 0.274)	-0.175	(-0.727, 0.376)	0.033	(-0.443, 0.508)	-0.353	(-0.992, 0.285)	-0.138	(-0.678, 0.401)
Dad job manual/farmer	-0.199	(-0.679, 0.280)	-0.503	(-1.120, 0.114)	-0.292	(-0.818, 0.233)	-0.134	(-0.582, 0.313)	-0.481	(-1.083, 0.121)	-0.260	(-0.771, 0.252)
Dad job military	-0.178	(-0.720, 0.364)	-0.451	(-1.116, 0.213)	-0.281	(-0.854, 0.292)	-0.074	(-0.564, 0.417)	-0.421	(-1.065, 0.223)	-0.236	(-0.786, 0.314)
Household size	0.001	(-0.033, 0.035)	-0.010	(-0.043, 0.024)	-0.005	(-0.037, 0.027)	0.002	(-0.029, 0.034)	-0.011	(-0.044, 0.023)	-0.007	(-0.038, 0.024)
School characteristics												
Small (1-400)	-0.005	(-0.434, 0.424)	-0.021	(-0.306, 0.263)	-0.018	(-0.282, 0.246)	-0.015	(-0.440, 0.410)	-0.023	(-0.302, 0.256)	-0.020	(-0.288, 0.248)
Large (1001-4000)	-0.087	(-0.245, 0.072)	-0.041	(-0.146, 0.065)	-0.033	(-0.131, 0.065)	-0.080	(-0.227, 0.066)	-0.049	(-0.147, 0.049)	-0.045	(-0.137, 0.046)
Urban	-0.059	(-0.227, 0.109)	-0.035	(-0.156, 0.087)	-0.034	(-0.149, 0.082)	-0.049	(-0.209, 0.112)	-0.033	(-0.148, 0.081)	-0.035	(-0.147, 0.078)
Rural	0.081	(-0.154, 0.316)	0.092	(-0.057, 0.241)	0.090	(-0.037, 0.217)	0.115	(-0.085, 0.315)	0.083	(-0.046, 0.211)	0.076	(-0.043, 0.195)
West	0.198	(-0.077, 0.472)	0.121	(-0.070, 0.313)	0.114	(-0.051, 0.278)	0.169	(-0.059, 0.398)	0.113	(-0.051, 0.276)	0.109	(-0.046, 0.264)
Midwest	0.105	(-0.149, 0.359)	0.060	(-0.123, 0.243)	0.063	(-0.073, 0.198)	0.054	(-0.148, 0.257)	0.038	(-0.101, 0.178)	0.041	(-0.091, 0.174)
South	0.111	(-0.127, 0.348)	0.068	(-0.089, 0.226)	0.069	(-0.065, 0.203)	0.092	(-0.113, 0.296)	0.060	(-0.081, 0.201)	0.063	(-0.063, 0.190)
Private	0.004	(-0.337, 0.344)	-0.021	(-0.248, 0.205)	-0.028	(-0.239, 0.184)	-0.004	(-0.349, 0.341)	-0.028	(-0.251, 0.195)	-0.034	(-0.247, 0.179)
Constant	5.628*	(4.898, 6.357)	4.949*	(4.342, 5.556)	4.822*	(4.211, 5.433)	5.436*	(4.737, 6.135)	4.921*	(4.315, 5.526)	4.819*	(4.221, 5.417)
Selection	-0.307	(-1.072, 0.458)	-0.073	(-0.610, 0.464)	-0.094	(-0.314, 0.126)						

a: Dynamic peer structure for endogenous peer effects, fixed peer structure for contextual peer effects. b: Dynamic peer structure for both types of peer effects. c: Fixed peer structure for both types of peer effects. All are estimated by NLS. 95% confidence interval in parentheses, computed from 500 bootstraps. * denotes CI does not include 0.

Table B.21: Social Studies GPAs with both endogenous and contextual peer effects - Marginal effects

	(I).a				(I).b				(I).c			
	Average own effect		Average total effect		Average own effect		Average total effect		Average own effect		Average total effect	
Age	-0.185*	(-0.241, -0.147)	-0.194*	(-0.254, -0.134)	-0.152*	(-0.350, -0.266)	-0.308*	(-0.433, -0.183)	-0.152*	(-0.350, -0.266)	-0.308*	(-0.433, -0.183)
Male	-0.110*	(-0.123, -0.048)	-0.086*	(-0.147, -0.024)	-0.111*	(-0.087, -0.013)	-0.050	(-0.141, 0.042)	-0.111*	(-0.087, -0.013)	-0.050	(-0.141, 0.042)
Grade 10	0.124	(-0.025, 0.244)	0.109	(-0.033, 0.251)	0.111*	(0.090, 0.340)	0.215*	(0.032, 0.398)	0.111*	(0.090, 0.340)	0.215*	(0.032, 0.398)
Grade 11	0.302*	(0.138, 0.466)	0.302*	(0.120, 0.484)	0.281*	(0.382, 0.675)	0.529*	(0.236, 0.821)	0.281*	(0.382, 0.675)	0.529*	(0.236, 0.821)
Grade 12	0.678*	(0.466, 0.836)	0.651*	(0.406, 0.896)	0.578*	(0.773, 1.116)	0.945*	(0.541, 1.348)	0.578*	(0.773, 1.116)	0.945*	(0.541, 1.348)
Hispanic	-0.103*	(-0.286, -0.160)	-0.223*	(-0.384, -0.062)	-0.104*	(-0.294, -0.171)	-0.233*	(-0.439, -0.026)	-0.104*	(-0.294, -0.171)	-0.233*	(-0.439, -0.026)
Black	-0.168*	(-0.337, -0.177)	-0.257*	(-0.372, -0.143)	-0.174*	(-0.365, -0.188)	-0.276*	(-0.419, -0.133)	-0.174*	(-0.365, -0.188)	-0.276*	(-0.419, -0.133)
Asian	0.179*	(0.168, 0.307)	0.237*	(0.068, 0.407)	0.190*	(0.175, 0.312)	0.243*	(0.021, 0.465)	0.190*	(0.175, 0.312)	0.243*	(0.021, 0.465)
Other races	-0.028	(-0.094, 0.014)	-0.040	(-0.158, 0.078)	-0.033*	(-0.130, -0.021)	-0.076	(-0.234, 0.083)	-0.033*	(-0.130, -0.021)	-0.076	(-0.234, 0.083)
Years in school	0.014	(-0.007, 0.039)	0.016	(-0.043, 0.074)	0.016*	(0.002, 0.046)	0.024	(-0.048, 0.095)	0.016*	(0.002, 0.046)	0.024	(-0.048, 0.095)
Live w mother	0.153*	(0.244, 0.373)	0.309*	(0.160, 0.457)	0.161*	(0.223, 0.350)	0.286*	(0.060, 0.513)	0.161*	(0.223, 0.350)	0.286*	(0.060, 0.513)
Live w father	0.034	(-0.024, 0.447)	0.211	(-0.218, 0.640)	0.014*	(0.383, 0.856)	0.620	(-0.288, 1.527)	0.014*	(0.383, 0.856)	0.620	(-0.288, 1.527)
Health	-0.118*	(-0.218, -0.190)	-0.204*	(-0.244, -0.165)	-0.116*	(-0.258, -0.230)	-0.244*	(-0.303, -0.185)	-0.116*	(-0.258, -0.230)	-0.244*	(-0.303, -0.185)
Mom edu not know	-0.031	(-0.068, 0.028)	-0.020	(-0.128, 0.088)	-0.030*	(-0.103, -0.006)	-0.054	(-0.224, 0.116)	-0.030*	(-0.103, -0.006)	-0.054	(-0.224, 0.116)
Mom edu less HS	-0.119*	(-0.163, -0.057)	-0.110*	(-0.217, -0.003)	-0.123*	(-0.116, -0.017)	-0.067	(-0.223, 0.090)	-0.123*	(-0.116, -0.017)	-0.067	(-0.223, 0.090)
Mom edu college	0.109*	(0.147, 0.208)	0.177*	(0.093, 0.262)	0.111*	(0.150, 0.210)	0.180*	(0.067, 0.293)	0.111*	(0.150, 0.210)	0.180*	(0.067, 0.293)
Dad edu not know	-0.132*	(-0.225, -0.083)	-0.154	(-0.313, 0.006)	-0.140*	(-0.231, -0.101)	-0.166	(-0.368, 0.036)	-0.140*	(-0.231, -0.101)	-0.166	(-0.368, 0.036)
Dad edu less HS	-0.061*	(-0.218, -0.093)	-0.156	(-0.322, 0.011)	-0.061*	(-0.266, -0.148)	-0.207	(-0.425, 0.010)	-0.061*	(-0.266, -0.148)	-0.207	(-0.425, 0.010)
Dad edu college	0.143*	(0.206, 0.289)	0.247*	(0.147, 0.348)	0.147*	(0.253, 0.334)	0.293*	(0.156, 0.431)	0.147*	(0.253, 0.334)	0.293*	(0.156, 0.431)
Mom job other	-0.141*	(-0.318, -0.174)	-0.246*	(-0.371, -0.121)	-0.147*	(-0.337, -0.194)	-0.265*	(-0.444, -0.087)	-0.147*	(-0.337, -0.194)	-0.265*	(-0.444, -0.087)
Mom job professionals	0.002	(-0.011, 0.084)	0.037	(-0.080, 0.153)	-0.000	(-0.001, 0.092)	0.045	(-0.120, 0.211)	-0.000	(-0.001, 0.092)	0.045	(-0.120, 0.211)
Mom job office/sales	-0.046*	(-0.116, -0.013)	-0.065	(-0.178, 0.049)	-0.048*	(-0.162, -0.059)	-0.111	(-0.273, 0.052)	-0.048*	(-0.162, -0.059)	-0.111	(-0.273, 0.052)
Mom job manual/farmer	-0.069*	(-0.208, -0.093)	-0.150	(-0.307, 0.006)	-0.073*	(-0.208, -0.097)	-0.153	(-0.376, 0.071)	-0.073*	(-0.208, -0.097)	-0.153	(-0.376, 0.071)
Mom job military	-0.127*	(-0.384, -0.087)	-0.235	(-0.623, 0.153)	-0.133*	(-0.465, -0.178)	-0.321	(-0.854, 0.211)	-0.133*	(-0.465, -0.178)	-0.321	(-0.854, 0.211)
Dad job other	-0.077*	(-0.472, -0.014)	-0.243	(-0.683, 0.197)	-0.055*	(-0.860, -0.396)	-0.628	(-1.531, 0.275)	-0.055*	(-0.860, -0.396)	-0.628	(-1.531, 0.275)
Dad job professionals	0.059	(-0.283, 0.212)	-0.036	(-0.480, 0.409)	0.082*	(-0.705, -0.212)	-0.459	(-1.351, 0.434)	0.082*	(-0.705, -0.212)	-0.459	(-1.351, 0.434)
Dad job office/sales	0.078	(-0.222, 0.290)	0.034	(-0.443, 0.511)	0.104*	(-0.576, -0.059)	-0.318	(-1.266, 0.631)	0.104*	(-0.576, -0.059)	-0.318	(-1.266, 0.631)
Dad job manual/farmer	-0.017	(-0.412, 0.064)	-0.174	(-0.584, 0.237)	0.007*	(-0.853, -0.376)	-0.615	(-1.476, 0.246)	0.007*	(-0.853, -0.376)	-0.615	(-1.476, 0.246)
Dad job military	0.064	(-0.325, 0.179)	-0.073	(-0.548, 0.402)	0.082*	(-0.694, -0.182)	-0.438	(-1.400, 0.524)	0.082*	(-0.694, -0.182)	-0.438	(-1.400, 0.524)
Household size	-0.005	(-0.020, 0.012)	-0.004	(-0.042, 0.033)	-0.003*	(-0.031, -0.001)	-0.016	(-0.076, 0.044)	-0.003*	(-0.031, -0.001)	-0.016	(-0.076, 0.044)

Marginal effects for (I) in Table B.20. Marginal effects for students selecting the subject only. Marginal effects average within in each subnetwork then average over subnetworks. a: Dynamic peer structure for endogenous peer effects, fixed peer structure for contextual peer effects. b: Dynamic peer structure for both types of peer effects. c: Fixed peer structure for both types of peer effects. 95% confidence interval in parentheses, computed from 500 bootstraps. * denotes CI does not include 0.

B.3 Network structure of contextual peer effects

In this section, we want to examine the results of Table 3.4 heuristically via simulation. We want to see the way estimates could potentially change based on different network structure assumptions on different types of peer effects.

We use the similar data to the Monte Carlo section of Chapter 2: friendship nomination in AddHealth and one covariate, which is health. However, in this section, we focus on the outcome (second) stage only and treat selection as random (at around 0.8). We consider three DGPS:

1. DGP(a): Endogenous peer effects are realized among selecting peers (peers who select into the outcome stage) while contextual peer effects are from all original peers.
2. DGP(b): Both endogenous peer effects and contextual peer effects come from selecting peers only.
3. DGP(c): Both endogenous peer effects and contextual peer effects come from original peers.

We estimate with three different NLS estimators in which the difference is in the network structure assumptions. Estimator (A) assumes the network structure of GDP(a) and therefore is correct under DGP(a). Similarly, (B) is correct under DPG(b), and (C) under DGP(c). We focus on the bias of the endogenous peer effects only, and under different strength of this endogenous peer effects, 0.05 and 0.5.

Under DGP(a), the bias of λ by (B) and (C) are in general small, and the direction of bias is unclear. Under DGP(b), the bias of (A) is much larger than (C). However, (A) in general overestimates λ rather than underestimate it. Under DGP(c), (A) again produces large bias but now with underestimation rather than overestimation. (B) does produce some underestimation as well.

Table B.22: Simulation results under DGP(a):
Different network structures for two types of peer effects

			ng = 20		ng = 40		ng = 80		ng = 160	
			Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Lambda: 0.05										
(A)	λ	0.050	-2e-03	[2e-03]	3e-03	[8e-04]	-6e-05	[4e-04]	6e-04	[2e-04]
	Cons	-0.423	-0.007	[1e-02]	-0.004	[6e-03]	-0.002	[3e-03]	0.000	[2e-03]
	β^{own}	-1.774	0.001	[1e-03]	0.001	[8e-04]	0.001	[4e-04]	-0.000	[2e-04]
	$\beta^{context}$	-0.073	-0.005	[7e-03]	0.005	[3e-03]	0.000	[2e-03]	0.002	[9e-04]
(B)	λ	0.050	-2e-03	[4e-02]	5e-03	[2e-02]	6e-04	[1e-02]	4e-03	[5e-03]
	Cons	-0.423	-0.022	[1e-02]	-0.018	[6e-03]	-0.019	[3e-03]	-0.016	[2e-03]
	β^{own}	-1.774	-0.002	[1e-03]	-0.001	[8e-04]	-0.001	[4e-04]	-0.002	[2e-04]
	$\beta^{context}$	-0.073	0.007	[2e-01]	0.020	[8e-02]	0.012	[4e-02]	0.018	[2e-02]
(C)	λ	0.050	4e-03	[5e-02]	2e-02	[2e-02]	1e-02	[1e-02]	1e-02	[6e-03]
	Cons	-0.423	-0.006	[1e-02]	-0.004	[6e-03]	-0.003	[3e-03]	-0.001	[2e-03]
	β^{own}	-1.774	0.000	[2e-03]	0.001	[8e-04]	0.001	[4e-04]	0.000	[2e-04]
	$\beta^{context}$	-0.073	0.014	[2e-01]	0.039	[1e-01]	0.035	[5e-02]	0.034	[3e-02]
Lambda: 0.5										
(A)	λ	0.500	8e-04	[4e-04]	1e-03	[2e-04]	-2e-04	[1e-04]	-2e-04	[6e-05]
	Cons	-0.423	0.001	[2e-02]	-0.003	[9e-03]	-0.002	[4e-03]	-0.001	[2e-03]
	β^{own}	-1.774	-0.000	[2e-03]	0.001	[1e-03]	0.001	[5e-04]	-0.000	[3e-04]
	$\beta^{context}$	-0.073	0.000	[6e-03]	0.004	[3e-03]	-0.000	[1e-03]	-0.000	[8e-04]
(B)	λ	0.500	3e-03	[1e-03]	1e-03	[6e-04]	-3e-04	[3e-04]	-3e-04	[1e-04]
	Cons	-0.423	-0.015	[2e-02]	-0.019	[8e-03]	-0.019	[4e-03]	-0.018	[2e-03]
	β^{own}	-1.774	-0.003	[2e-03]	-0.001	[1e-03]	-0.001	[5e-04]	-0.001	[2e-04]
	$\beta^{context}$	-0.073	0.019	[1e-02]	0.013	[7e-03]	0.009	[3e-03]	0.010	[2e-03]
(C)	λ	0.500	5e-03	[2e-03]	6e-03	[9e-04]	4e-03	[4e-04]	5e-03	[2e-04]
	Cons	-0.423	-0.049	[2e-02]	-0.052	[1e-02]	-0.054	[8e-03]	-0.051	[5e-03]
	β^{own}	-1.774	0.009	[2e-03]	0.011	[1e-03]	0.012	[7e-04]	0.010	[4e-04]
	$\beta^{context}$	-0.073	0.086	[3e-02]	0.089	[2e-02]	0.085	[1e-02]	0.087	[1e-02]

Table B.23: Simulation results under DGP(b):
Same dynamic network structure for both types of peer effects

			ng = 20		ng = 40		ng = 80		ng = 160	
			Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Lambda: 0.05										
(A)	λ	0.050	3e-02	[3e-03]	4e-02	[2e-03]	3e-02	[2e-03]	3e-02	[1e-03]
	Cons	-0.423	-0.001	[1e-02]	0.002	[6e-03]	0.004	[3e-03]	0.006	[2e-03]
	β^{own}	-1.774	0.001	[1e-03]	0.001	[8e-04]	0.000	[4e-04]	-0.001	[2e-04]
	$\beta^{context}$	-0.073	0.066	[1e-02]	0.075	[9e-03]	0.071	[7e-03]	0.072	[6e-03]
(B)	λ	0.050	-5e-03	[4e-02]	2e-03	[2e-02]	-3e-03	[1e-02]	4e-05	[5e-03]
	Cons	-0.423	-0.004	[1e-02]	-0.001	[6e-03]	-0.001	[3e-03]	0.001	[2e-03]
	β^{own}	-1.774	-0.001	[1e-03]	0.001	[8e-04]	0.000	[4e-04]	-0.000	[2e-04]
	$\beta^{context}$	-0.073	-0.008	[2e-01]	0.003	[8e-02]	-0.006	[4e-02]	0.000	[2e-02]
(C)	λ	0.050	2e-02	[5e-02]	3e-02	[2e-02]	3e-02	[1e-02]	3e-02	[7e-03]
	Cons	-0.423	-0.006	[1e-02]	-0.004	[6e-03]	-0.004	[3e-03]	-0.002	[2e-03]
	β^{own}	-1.774	0.001	[2e-03]	0.002	[8e-04]	0.002	[4e-04]	0.001	[2e-04]
	$\beta^{context}$	-0.073	0.042	[2e-01]	0.070	[1e-01]	0.066	[6e-02]	0.066	[3e-02]
Lambda: 0.5										
(A)	λ	0.500	1e-02	[6e-04]	1e-02	[4e-04]	1e-02	[2e-04]	1e-02	[2e-04]
	Cons	-0.423	0.011	[2e-02]	0.007	[8e-03]	0.007	[4e-03]	0.009	[2e-03]
	β^{own}	-1.774	-0.005	[2e-03]	-0.004	[1e-03]	-0.004	[5e-04]	-0.005	[3e-04]
	$\beta^{context}$	-0.073	0.043	[8e-03]	0.048	[5e-03]	0.043	[3e-03]	0.043	[3e-03]
(B)	λ	0.500	3e-03	[1e-03]	8e-04	[6e-04]	-7e-04	[3e-04]	-6e-04	[1e-04]
	Cons	-0.423	0.002	[2e-02]	-0.002	[8e-03]	-0.002	[4e-03]	-0.001	[2e-03]
	β^{own}	-1.774	-0.002	[2e-03]	0.001	[1e-03]	0.001	[5e-04]	0.000	[2e-04]
	$\beta^{context}$	-0.073	0.008	[1e-02]	0.002	[7e-03]	-0.002	[3e-03]	-0.002	[2e-03]
(C)	λ	0.500	7e-03	[2e-03]	8e-03	[9e-04]	6e-03	[5e-04]	7e-03	[3e-04]
	Cons	-0.423	-0.049	[2e-02]	-0.053	[1e-02]	-0.055	[8e-03]	-0.052	[5e-03]
	β^{own}	-1.774	0.009	[2e-03]	0.011	[1e-03]	0.012	[7e-04]	0.010	[4e-04]
	$\beta^{context}$	-0.073	0.097	[3e-02]	0.099	[2e-02]	0.096	[1e-02]	0.097	[1e-02]

Table B.24: Simulation results under DGP(c):
Same fixed network structure for both types of peer effects

			ng = 20		ng = 40		ng = 80		ng = 160	
			Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
Lambda: 0.05										
(A)	λ	0.050	-5e-02	[4e-03]	-5e-02	[3e-03]	-5e-02	[3e-03]	-5e-02	[2e-03]
	Cons	-0.423	-0.014	[1e-02]	-0.011	[6e-03]	-0.009	[3e-03]	-0.007	[2e-03]
	β^{own}	-1.774	0.001	[1e-03]	0.002	[8e-04]	0.001	[4e-04]	-0.000	[2e-04]
	$\beta^{context}$	-0.073	-0.103	[2e-02]	-0.093	[1e-02]	-0.098	[1e-02]	-0.096	[1e-02]
(B)	λ	0.050	-6e-03	[4e-02]	3e-03	[2e-02]	-2e-03	[1e-02]	2e-03	[5e-03]
	Cons	-0.423	-0.047	[1e-02]	-0.044	[8e-03]	-0.044	[5e-03]	-0.041	[3e-03]
	β^{own}	-1.774	-0.004	[2e-03]	-0.003	[8e-04]	-0.003	[4e-04]	-0.004	[2e-04]
	$\beta^{context}$	-0.073	0.012	[2e-01]	0.028	[9e-02]	0.021	[4e-02]	0.028	[2e-02]
(C)	λ	0.050	-9e-03	[5e-02]	3e-03	[2e-02]	2e-04	[1e-02]	-8e-04	[6e-03]
	Cons	-0.423	-0.004	[1e-02]	-0.002	[6e-03]	-0.001	[3e-03]	0.001	[2e-03]
	β^{own}	-1.774	-0.000	[1e-03]	0.001	[8e-04]	0.000	[4e-04]	-0.000	[2e-04]
	$\beta^{context}$	-0.073	-0.018	[2e-01]	0.006	[9e-02]	0.001	[5e-02]	-0.001	[2e-02]
Lambda: 0.5										
(A)	λ	0.500	-3e-01	[9e-02]	-3e-01	[9e-02]	-3e-01	[9e-02]	-3e-01	[9e-02]
	Cons	-0.423	-0.206	[6e-02]	-0.208	[5e-02]	-0.204	[5e-02]	-0.202	[4e-02]
	β^{own}	-1.774	0.021	[3e-03]	0.022	[2e-03]	0.021	[1e-03]	0.020	[7e-04]
	$\beta^{context}$	-0.073	-1.037	[1e+00]	-1.033	[1e+00]	-1.041	[1e+00]	-1.041	[1e+00]
(B)	λ	0.500	-6e-02	[5e-03]	-6e-02	[4e-03]	-6e-02	[4e-03]	-6e-02	[4e-03]
	Cons	-0.423	-0.410	[2e-01]	-0.413	[2e-01]	-0.413	[2e-01]	-0.413	[2e-01]
	β^{own}	-1.774	-0.042	[4e-03]	-0.040	[3e-03]	-0.042	[2e-03]	-0.042	[2e-03]
	$\beta^{context}$	-0.073	0.030	[2e-02]	0.033	[9e-03]	0.027	[5e-03]	0.027	[3e-03]
(C)	λ	0.500	1e-03	[1e-03]	1e-03	[6e-04]	-3e-04	[3e-04]	-4e-04	[1e-04]
	Cons	-0.423	0.002	[1e-02]	-0.003	[8e-03]	-0.002	[4e-03]	-0.000	[2e-03]
	β^{own}	-1.774	-0.001	[2e-03]	0.001	[1e-03]	0.001	[4e-04]	-0.000	[2e-04]
	$\beta^{context}$	-0.073	0.002	[1e-02]	0.005	[7e-03]	-0.001	[4e-03]	-0.001	[2e-03]

Bibliography

Angrist, J. D. and Lang, K. (2004). Does school integration generate peer effects? Evidence from Boston's METCO program. *American Economic Review*, 94(5):1613–1634.

Badev, A. (2017). Discrete games in endogenous networks: Theory and policy. Revise and submitted at Econometrica.

Bhattacharyya, S. and Bickel, P. J. (2015). Subsampling bootstrap of count features of networks. *Annals of Statistics*, 43(6):2384–2411.

Blume, L. E., Brock, W. A., Durlauf, S. N., and Jayaraman, R. (2015). Linear social interactions models. *Journal of Political Economy*, 123(2):444–496.

Boucher, V. (2016). Conformism and self-selection in social networks. *Journal of Public Economics*, 136:30–44.

Boucher, V. and Fortin, B. (2016). Some challenges in the empirics of the effects of networks. In Bramoullé, Y., Galeotti, A., and Rogeres, B., editors, *The Oxford Handbook of the Economics of Networks*. Oxford University Press, Oxford.

Bramoullé, Y., Djebbari, H., and Fortin, B. (2009). Identification of peer effects through social networks. *Journal of Econometrics*, 150(1):41–55.

Carrell, S. E., Fullerton, R. L., and West, J. E. (2009). Does your cohort matter? Measuring peer effects in college achievement. *Journal of Labor Economics*, 27(3):439–464.

Carrell, S. E., Sacerdote, B. I., and West, J. E. (2013). From natural variation to optimal policy? The importance of endogenous peer group formation. *Econometrica*, 81(3):855–882.

Chandrasekhar, A. G. and Lewis, R. (2016). Econometrics of sampled networks. Working paper.

Federman, M. (2007). State graduation requirements, high school course taking, and choosing a technical college major. *B.E. Journal of Economic Analysis and Policy*, 7(1):1–34.

Flores-Lagunes, A. and Schnier, K. E. (2012). Estimation of sample selection models with spatial dependence. *Journal of Applied Econometrics*, 27:173–204.

Fruehwirth, J. C. (2014). Can achievement peer effect estimates inform policy? A view from inside the black box. *The Review of Economics and Statistics*, 96(3):514–523.

Gel, Y. R., Lyubchich, V., and Ramirez, L. L. R. (2017). Bootstrap quantification of estimation uncertainties in network degree distributions. *Scientific Reports*, 7(1):5807.

Goldsmith-Pinkham, P. and Imbens, G. W. (2013). Social networks and the identification of peer effects. *Journal of Business & Economics Statistics*, 31(3):253–264.

Goodman, J. (2019). The labor of division: Returns to compulsory high school math coursework. *Journal of Labor Economics*, 37(4):1141–1182.

Hao, Z. and Cowan, B. W. (2019). The effects of graduation requirements on risky health behaviors of high school students. *American Journal of Health Economics*, 5(1):97–125.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1):153–161.

Hoshino, T. (2017). Two-step estimation of incomplete information social interaction models with sample selection. *Journal of Business and Economic Statistics*.

Hsieh, C.-S., Kho, S. I. M., Kovaik, J., and Logan, T. (2019). Non-randomly sampled networks: Biases and corrections. NBER Working paper 25270.

Hsieh, C.-S. and Lee, L. (2016). A social interactions model with endogenous friendship formation and selectivity. *Journal of Applied Econometrics*, 31:301–319.

Hsieh, C.-S. and van Kipperluis, H. (2018). Smoking initiation: Peers and personality. *Quantitative Economics*, 9:825–863.

Imberman, S. A., Kugler, A. D., and Sacerdote, B. I. (2012). Katrina's children: Evidence on the structure of peer effects from hurricane evacuees. *American Economic Review*, 102(5):2048–2082.

Jin, F. and Lee, L. (2015). On the bootstrap for moran's i test for spatial dependence. *Journal of Econometrics*, 184:295–314.

Kelejian, H. H. and Prucha, I. R. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Estate Finance and Economics*, 17(1):99–121.

Lazarsfeld, P. F. and Merton, R. K. (1954). Friendship as social process: A substantive and methodological analysis. In Berger, M., editor, *Freedom and Control in Modern Society*, pages 18–66. Van Nostrand, New York.

Lee, L., Li, J., and Lin, X. (2014). Binary choice models with social network under heterogeneous rational expectations. *The Review of Economics and Statistics*, 96(3):402–417.

Lee, S. H., Kim, P.-J., and Jeong, H. (2006). Statistical properties of sampled networks. *Physical Review E*, 73:016102.

Levine, P. and Zimmerman, D. (1995). The benefit of additional high-school math and science classes for young men and women. *Journal of Business and Economic Statistics*, 13(2):137–149.

Lin, X. (2010). Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables. *Journal of Labor Economics*, 28(4):825–860.

Liu, X., Patacchini, E., and Rainone, E. (2017). Peer effects in bedtime decisions among adolescents: A social network model with sampled data. *The Econometrics Journal*, 20:S103–S125.

Liu, X., Patacchini, E., and Zenou, Y. (2014). Endogenous peer effects: Local aggregate or local average? *Journal of Economic Behavior and Organization*, 103:39–59.

Mangrum, D. (2019). Personal finance education mandates and student loan repayment. SSRN Electronic Journal working paper.

Manski, C. (1993). Identificiation of endogenous social effects: The reflection problem. *Review of Economic Studies*, 60(3):531–542.

McMillen, D. P. (1995). Selection bias in spatial econometrics models. *Journal of Regional Science*, 35(3):417–436.

McPherson, M., Smith-Lovin, L., , and Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Reviews of Sociology*, 27:415–444.

Patacchini, E., Rainone, E., and Zenou, Y. (2017). Heterogeneous peer effects in education. *Journal of Economic Behavior and Organization*, 134:190–227.

Pinkse, J. and Slade, M. E. (1998). Contracting in space: An application of spatial statistics to discrete-choice models. *Journal of Econometrics*, 85(1):125–154.

Rabovic, R. and Cizek, P. (2016). Estimation of spatial sample selection models: A partial maximum likelihood approach. Working paper.

Sacerdote, B. I. (2001). Peer effects with random assignment: Results for Dartmouth roommates. *Quarterly Journal of Economics*, 116(2):681–704.

Schiller, K. S. and Muller, C. (2003). Raising the bar and equity? effects of state high school graduation requirements and accountability policies on students' mathematics course taking. *Educational Evaluation and Policy Analysi*, 25(3):299–318.

Snijders, T., Steglich, C., and michael Schweinberger (2007). Modeling the coevolution of networks and behavior. In van Montfort, K., Oud, J., and Satorra, A., editors, *Longitudinal models in behavioral and related sciences*, chapter 3. Routledge, New York.

Teitelbaum, P. (2003). The influence of high school graduation requirement policies in mathematics and science on student course-taking patterns and achievement. *Educational Evaluation and Policy Analysis*, 23(1):31–57.

Thompson, M. E., Ramirez, L. L. R., Lyubchich, V., and Gel, Y. R. (2016). Using the bootstrap for statistical inference on random graphs. *The Canadian Journal of Statistics*, 44(1):3–24.

Wang, W. and Lee, L. (2013). Estimation of spatial autoregressive models with randomly missing data in the dependent variable. *The Econometrics Journal*, 16:73–102.

Xu, H. (2018). Social interactions in large networks: A game theoretic approach. *International Economic Review*, 59(1):257–284.

Yang, C. and Lee, L. (2017). Social interactions under incomplete information with heterogeneous expectations. *Journal of Econometrics*, 198:65–83.