NBER WORKING PAPER SERIES

SOCIAL INTERACTIONS, MECHANISMS, AND EQUILIBRIUM: EVIDENCE FROM A MODEL OF STUDY TIME AND ACADEMIC ACHIEVEMENT

Timothy Conley
Nirav Mehta
Ralph Stinebrickner
Todd Stinebrickner

Working Paper 21418
http://www.nber.org/papers/w21418

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2015, Revised February 2018

All authors are affiliated with the University of Western Ontario. This work has been made possible by generous funding provided by NSF, SSHRC, the Mellon Foundation, and the Spencer Foundation. George Orlov and Zinaida Foltin provided research assistance. We have benefited from comments made during seminars at McMaster University, Cornell PAM, and UCLA, and the Banff Empirical Microeconomics Workshop and CESifo Area Conference on the Economics of Education. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Social Interactions, Mechanisms, and Equilibrium: Evidence from a Model of Study Time and Academic Achievement
Timothy Conley, Nirav Mehta, Ralph Stinebrickner, and Todd Stinebrickner
NBER Working Paper No. 21418
July 2015, Revised February 2018
JEL No. H0,I20,J0

ABSTRACT

We develop and estimate a model of student study time on a social network. The model is designed to exploit unique data collected in the Berea Panel Study. Study time data allow us to quantify an intuitive mechanism for academic social interactions: own study time may depend on friend study time in a heterogeneous manner. Social network data allow us to embed study time and resulting academic achievement in an estimable equilibrium framework. We develop a specification test that exploits the equilibrium nature of social interactions and use it to show that novel study propensity measures mitigate econometric endogeneity concerns.

Timothy Conley
Department of Economics
Social Science Centre, Room 4071
London, Ontario, Canada, N6A 5C2
tconley3@uwo.ca

Nirav Mehta
University of Western Ontario
nirav.mehta@uwo.ca

Ralph Stinebrickner
Berea College
Berea, KY 40403
and University of Western Ontario
ralph_stinebrickner@berea.edu

Todd Stinebrickner
Department of Economics
Western University
London, Ontario, N6A 5C2
CANADA
and NBER
trstineb@uwo.ca

1 Introduction

Peer effects are widely believed to be important for determining academic achievement. Much of the existing research in this context has focused on establishing a causal link between peer characteristics and academic outcomes, in an effort to provide evidence about whether peers matter. However, though crucial for policymaking, direct evidence on the mechanisms generating peer effects is limited. In this paper we exploit unique data on college students from the Berea Panel Study (BPS) to study peer effects in an academic setting. We focus on what is likely the most relevant set of peers in our higher education context, a student's friends.

The goal of this paper is to better understand how peer effects are generated. One step is to provide direct evidence about a mechanism underlying peer influences in our context. This is in the spirit of Manski (2000), who stresses that, in order to understand relationships between own and peer outcomes, it is important to clearly define mechanisms and to obtain direct evidence about their relevance. Motivated by recent research hypothesizing that student effort is likely to be an input that is readily influenced by peers in the short run (Stinebrickner and Stinebrickner (2006), Calvó-Armengol et al. (2009), Cooley Fruehwirth (2013), and De Giorgi and Pellizzari (2014)), we focus on study time as an explicit mechanism through which peer effects could arise in college.[1]

Another step is to better understand the role social networks play in the propagation of peer effects. Not only may student i's study time be influenced by i's peers, but i's peers' study time may be influenced by i. Moreover, these types of feedback effects could work indirectly through students in the social network who are not directly connected to student i. We focus on how the distribution of feedback effects depends on three interrelated components. First, the graph describing links in a social network, which we refer to as the "network structure", may be important in and of itself (Calvó-Armengol et al. (2009), Jackson and Yariv (2011)). Second, students with different characteristics may differ in how much they are affected by their peers (Sacerdote (2011)). Third, students may form links based on these characteristics. In particular, students may link to others with similar characteristics, i.e., the network may exhibit "homophily." Together, the network structure and the specific manner in which heterogeneous students are arranged on the network may determine how changes in behavior propagate throughout the network and affect equilibrium outcomes.

To take these important next steps, we estimate an equilibrium model of study time

[1] For a non-education (financial) example of research that is interested in understanding why peer effects exist, see Bursztyn et al. (2014). Richards-Shubik (2015) separates supply and demand mechanisms in a model of sexual initiation.

choice and resulting grade determination, given a social network. Estimating such a model entails substantial data challenges. First, we need student-level data on study time. Unfortunately, because collecting reliable time-use information is very difficult in annual surveys, available data sources typically do not contain this type of information. Second, equilibrium outcomes depend on the entire social network, necessitating data characterizing the full set of peer connections as well as data on characteristics that likely determine study time choices. Among existing sources of social network data, perhaps only one, the National Longitudinal Survey of Adolescent Health (Add-Health), could potentially provide a full view of a social network in an educational setting where academic outcomes and student characteristics are also observed. Unfortunately, because the Add-Health dataset has a primary focus on adolescent health and risk-related behaviors, it does not contain information about time spent studying. Thus, to the best of our knowledge, there is no other data source that is able to both fully characterize a social network of students and provide direct evidence about a central input in the grade production function that has been hypothesized to generate social interactions.

Our project is made possible by unique data from the Berea Panel Study (BPS), which were collected specifically to overcome these current data limitations. The BPS surveyed full cohorts of students at Berea College, which allows us to characterize the entire social network. The BPS is also unique in its high frequency of contact with students each year, allowing the collection of eight time-use diaries, which allow us to characterize study time, and the measurement of friendships in each semester, which we use to define peers. We combine these survey data with administrative data that include pre-college characteristics and college grades.

We develop our model to exploit these unique data. The social network is known at the beginning of a period. Subsequently, all students in the social network simultaneously choose their study time to maximize their own achievement, net of studying costs. A student's studying cost depends on her own study time and friend study time, e.g., it may be more fun to study if your friends are studying (alternatively, students may conform to their friends). Cost functions are allowed to be heterogeneous across students. Achievement depends on a student's own study time and may also be heterogeneous across students, conditional on own study time.

The social interactions literature has paid close attention to the endogeneity problem that is present if there exist correlated unobserved variables, that is, unobserved information related to both peer group membership (in our context, friendship choices) and outcomes of interest (Manski (1993), Moffitt (2001), Epple and Romano (2011)). In our case, where we focus on a social interaction in study time choices, a relationship between friends' study times could arise because friends influence each other (peer effects) or because students with similar unobserved determinants of study time become friends (correlated unobserved variables). Institutional details, together with empirical checks we conduct, suggest that correlated shocks arising through, e.g., coursework and dormitories, are not the most salient type of correlated unobserved variables. The most relevant type of correlated unobserved variable would seem to be an unobserved individual characteristic, which could be thought of as a student's propensity to study.

We adopt a two-step approach for dealing with this endogeneity problem. First, we take advantage of a unique opportunity to directly measure students' propensities to study. Specifically, the day before freshman classes began, we collected information about how much a student actually studied in high school and how much the student expected to study in college. We find that both high school study time and expected college study time have strong correlations with study time in college and are also strongly related to friendship patterns in our data.

One cannot know *a priori* whether our study propensity data address the above endogeneity concern in a satisfactory manner, meaning we need some way to assess how well our data have measured typical correlated unobserved variables. Given the importance of this assessment, our second step is to develop a specification test based on our model. Our specification test is designed to detect unobserved determinants of study time, exploiting the fact that the equilibrium nature of social interactions implies that such unobserved determinants would generate cross-sectional dependence in residuals. Crucially, our test is useful even when unobserved determinants lead to inconsistently estimated parameters.

We estimate the model using data from two semesters. Under the baseline specification, in which we use our study propensity data to estimate the model, we find no evidence of the cross-sectional residual correlation described above. However, we do find significant cross-sectional residual correlations when we re-estimate the model excluding our study propensity data, i.e., using only measures of student characteristics that are typically available to researchers. This suggests that our specification test has the power to detect unobserved determinants of study time. Therefore, these findings provide evidence that our study propensity measures play an important role in addressing endogeneity concerns.

Our estimates provide strong evidence that friend study time has a substantial effect on one's own study time. We also find that one's own study time is an important determinant of one's own achievement. We estimate students to have different best response functions, i.e., they react differently to changes in friend study time. Hereafter, we will often refer to this as heterogeneity in *reactiveness*. This heterogeneity potentially has equilibrium implications, as it implies complementarities in students' choice of study time. We estimate that two students with 75th percentile reactiveness, when paired with each other, would study almost twice as much as would two students with 25th percentile reactiveness, when paired with each other.

The extent to which heterogeneity in reactiveness affects total production depends on the relationship between own and friend reactiveness. Therefore, it is also important to take into account the social network to understand social interactions.[2] We use our estimated model to perform two counterfactual exercises. First, we examine how the network structure, combined with homophilous sorting into friendships, affects the response to changes in friend study time. We exogenously increase (shock) the study time of each student and assess how study times and achievement change for other students in the social network. There is substantial heterogeneity in study time responses depending on which student is shocked, with larger impacts associated with more central students and students connected to more reactive peers. The specific manner in which students with different characteristics are arranged on the network is important for responses. This exercise also provides a natural framework for quantifying the importance of equilibrium interactions. On average, equilibrium responses produce a network-wide aggregate response that is 2.7 times larger than their partial equilibrium counterparts, which only consider a shock's effect on immediate neighbors.

Our framework allows us to provide further evidence about the importance of homophily in determining outcomes. As Golub and Jackson (2012) note, despite a large amount of work documenting the existence of homophily and a smaller literature examining its origins, the literature modeling the effect of homophily is in its infancy.[3] In our second counterfactual, we examine how achievement would differ if friend characteristics were identically distributed across students, instead of being strongly correlated with one's own characteristics, or homophilous, as in the data. On average, women, blacks, and students with above-median high school GPAs have high propensities to study and tend, in the data, to sort into friendships with students similar to themselves. Therefore, these groups tend to see declines in their friends' propensities to study in the counterfactual. However, these groups' losses are not offset by the gains of their complements. Intuitively, the estimated heterogeneity in best response functions means that total study time (and hence, achievement) is highest when students with high propensities to study are friends with others with high propensities to

[2]Kline and Tamer (2011) discuss the importance of distinguishing between estimates of technological parameters and the equilibrium effects of social interactions.

[3]Jackson (2008) provides a discussion of work documenting the existence of homophily; see Camargo et al. (2010) for a specific example. For theoretical models of homophily's origins see Currarini et al. (2009), Currarini et al. (2010), and Bramoullé et al. (2012). Badev (2013) allows for homophily in his empirical study of friendship formation and smoking behavior.

study, as is on average the case in the data. In contrast, there is a lack of such assortative matching in the counterfactual networks, meaning the economy does not take advantage of the game's supermodular structure.

The remainder of this paper is organized as follows. Section 2 discusses related literature. Section 3 contains a description of the BPS data. Section 4 presents our model. Section 5 presents our empirical specification. Section 6 develops our specification test and Section 7 discusses estimation results. Section 8 presents the results from our counterfactual exercises and Section 9 concludes.

2 Related Literature

Academic Peer Effects and Social Interactions Models There is an extensive literature on academic peer effects, which has been recently surveyed by Epple and Romano (2011) and Sacerdote (2011). As discussed in Sacerdote (2011), papers in this literature typically do not directly examine mechanisms through which peer effects are generated. Cooley Fruehwirth (2013), Calvó-Armengol et al. (2009), De Giorgi and Pellizzari (2014), and Tincani (2016) all stress the importance of equilibrium models of students' effort choices, but lack direct data on student effort. Cooley Fruehwirth (2013) and Calvó-Armengol et al. (2009) estimate parameters of their respective models, identifying effort through residual variation in peer outcomes. De Giorgi and Pellizzari (2014) and Tincani (2016) test the implications of different theoretical models of social interactions using student achievement data.

In terms of goals, perhaps the paper closest to ours is Calvó-Armengol et al. (2009). As in that paper, we develop a model that takes the social network as given to understand how effort choices, made on the social network, affect academic achievement. Calvó-Armengol et al. (2009), lacking direct input data, consider an environment in which a socially determined input choice is linked to network topology. This provides a behavioral foundation to the Katz-Bonacich centrality measure. Our contribution is the direct measurement of an input that likely affects achievement (study effort) and variables that likely are related to sorting into friendships and the choice of this input (study propensity measures, like high school study time and expected study time). These data are crucial for thoroughly investigating the mechanism considered in this paper. They allow us to examine how the input of interest (study effort) influences the output of interest (achievement). They also allow us to quantify how the input of interest is influenced by peers. Our theoretical model differs from that in Calvó-Armengol et al. (2009) in potentially important ways that are testable. We allow for heterogeneous best response functions, which our data on input choices allow us to identify. We also allow for nonlinearity in best response functions, which would break the connection between network topology and equilibrium outcomes required in Calvó-Armengol et al. (2009).

Our approach complements that of Calvó-Armengol et al. (2009) by allowing for a richer understanding of social interactions. In our framework, someone concerned about a student's low effort level may have an incentive to get the student to have more studious friends. Because we allow inputs to depend on student characteristics, there may be winners and losers from changes in peer group composition. Moreover, potentially heterogeneous reactiveness would allow total achievement to change, depending on the type of sorting in the baseline. Such questions could not be assessed using the framework of Calvó-Armengol et al. (2009), where counterfactuals are limited to changes in link structure.

There is a growing literature studying peer effects that has focused on modeling the formation of social networks, an important and notoriously difficult problem (see Christakis et al. (2010), Mele (2013), Badev (2013), de Paula et al. (2016), Sheng (2014), and Hsieh and Lee (2016)). We cannot study how the network would change in response to a policy because we do not model how friendships are formed. Therefore, in our counterfactuals, we examine fully-specified networks of interest, such as those in the data and randomly generated networks, in which student and friend characteristics are independent.

Specification Test The specification test we develop is informative about the presence of unobserved determinants of study time of the sort discussed in the introduction, even when parameter estimates are biased. Our specification test exploits the fact that, in equilibrium, all unobserved determinants of study time will typically enter all students' outcome equations. This type of error structure has precedent in a social interactions context (see, e.g., Calvó-Armengol et al. (2009) and Blume et al. (2015)).[4] Our contribution is that we develop a specification test designed to detect unobserved variables of interest and show how it can have the power to do so, even in the presence of inconsistently estimated parameters.

Goldsmith-Pinkham and Imbens (2013) posit a model of network formation and, within this model, derive a testable implication of endogenous network formation (see Boucher and Fortin (2016) for further discussion). In contrast, we do not test for a specific model of network formation. Rather, the goal of our specification test is to detect unobserved determinants of study time that we believe to be relevant to our context, taking as given the network. Therefore, we view our work as complementary to that of Goldsmith-Pinkham and Imbens (2013).

[4]There is a related literature on spatial autoregressive models; see Pinkse et al. (2002) and Lee (2004), for example.

3 Data

The BPS is a longitudinal survey that was designed by Todd Stinebrickner and Ralph Stinebrickner to provide detailed information about educational outcomes in college and labor market outcomes in the early post-college period. The BPS survey design involved collecting information about all students who entered Berea College in the fall of 2000 and the fall of 2001. Baseline surveys were conducted immediately before the start of first year classes and students were subsequently surveyed 10-12 times each year during school. As has been discussed in previous work that uses the BPS, caution is appropriate when considering exactly how results from the BPS would generalize to other specific institutions (e.g., Stinebrickner and Stinebrickner (2006, 2013)). At the same time, from an academic standpoint, Berea has much in common with many four-year colleges. It operates under a standard liberal arts curriculum and the students at Berea, which is in central Kentucky, are similar in academic quality to, for example, students at the University of Kentucky (Stinebrickner and Stinebrickner (2008b)).

Our study is made possible by three types of information that are available in the BPS. First, the BPS elicited each student's closest friends. Our analysis utilizes friendship observations from the end of the first semester and the end of the second semester. The survey question for the end of the first semester is shown in Appendix A.1. The survey question for the end of the second semester is identical (except for the date). Our friendship survey questions have a full-semester flavor to them, as they asked students to list the four people who had been their best friends that semester. Second, the BPS collected detailed time-use information eight times each year; for our sample, this was done using the twenty-four hour time diary shown in Appendix A.1. Finally, questions on the baseline survey reveal the number of hours that a student studied per week in high school and how much the student expects to study per week in college. We refer to these variables as our study propensity measures. The survey data are merged with detailed administrative data on race, sex, high school grade point average (GPA), college entrance exam scores, and college GPA in each semester.

This paper focuses on the freshman year for students in the 2001 entering cohort.[5] We focus on understanding grade outcomes during the freshmen year for two primary reasons. First, under the general liberal arts curriculum, students tend to have similar course loads

[5] We focus on this cohort because the survey contains more comprehensive time-use and friendship information for them. Information about time use was collected using time diaries for the 2001 cohort, while, for the 2000 cohort, this information was collected using questions that asked respondents to "think carefully about how much time was spent studying" in the last twenty-four hours. First semester friendship information was collected at the end of the first semester for the 2001 cohort, while, for the 2000 cohort, first semester friendship information was collected retrospectively during the second semester.

in their first year. Second, we are able to characterize the network most completely in the first year both because survey response rates are very high in the first year and because over 80% of friends reported by students in their freshman year are themselves freshmen.[6] These advantages tend to fade in subsequent years as friendships change (in part, due to dropout after the first year) and students' programs of study specialize.

3.1 Sample Construction

Our focus is on students who stayed in school for the full first year. There were a total of 331 students who fit this description. Our estimation sample consists of the 307 students (i.e., 93% of the 331) with friends in each semester. A student j is deemed to be a friend of student i if either i lists j as a friend or j lists i as a friend. This means that a student can have friends in a particular semester even if the student did not complete the friendship question in that semester. However, pooling the two semesters, we find that about 85% of the students in our final sample reported friendship information directly, via the friendship survey.

3.2 Descriptive Statistics

Tables 1-3 contain descriptive statistics for the sample. Table 1 shows descriptive statistics of student characteristics. The first row in each of the six panels shows overall descriptive statistics for the variable of interest described in the first column. Forty-four percent of students are male, 18% of students are black, the mean high school grade point average for the sample is 3.39, the mean combined score on the American College Test (ACT) is 23.26, and, on average, students studied 11.24 hours per week in high school and expect to study 24.96 hours per week in college. The subsequent rows in each panel show descriptive statistics for the variable of interest in the first column for different groups. For example, the third panel shows that, on average, males have lower high school grade point averages than females (3.24 vs. 3.51) and blacks have lower high school grade point averages than nonblacks (3.14 vs. 3.45). The fifth panel shows that blacks studied more, on average, in high school than other students (15.29 vs. 10.36).[7]

Table 2 shows descriptive statistics of outcomes during the first year. The first rows of panels 1 and 2, respectively, show that, on average, students study 3.49 hours per day in the

[6] Approximately 88% of all entering students in the 2001 cohort completed our baseline survey, and response rates remained high for the eleven subsequent surveys that were administered during the freshman year.

[7] The first two differences in means are significantly different at the 0.001 level. The averages of high school study time for blacks and nonblacks are significantly different at the 0.01 level.

first semester and 3.5 hours per day in the second semester.[8] The subsequent rows of the first two panels show that, on average, males study less than females, blacks study more than nonblacks, and students with above-median high school GPAs study more than students with below-median high school GPAs.[9] The first rows of panels 3 and 4, respectively, show that the average first semester GPA is 2.89 and the average second semester GPA is 2.93. The subsequent rows of the third and fourth panels show that males, blacks, and students with below-median high school GPAs all have lower average GPAs than their counterparts.[10]

As described at the beginning of this section, we define friendship as the union of reported links between two students that semester.[11] Table 3 summarizes friend data for those who have at least one friend in each semester, stratified by the same characteristics as in Table 1. The top panel shows that students have 3.3 friends on average. The mean masks considerable variation: the minimum number of friends is one, while the maximum number of friends is 10. The second and third panels show that male and black students (and, therefore, female and nonblack students) sort strongly towards students with the same characteristics. For example, 74% of the friends of male students are male, while only 18% of the friends of female students are male. Similarly, 69% of the friends of black students are black, while only 7% of the friends of nonblack students are black. The fourth and fifth panels show that male and black students have friends with lower incoming GPAs and lower combined ACT scores. The sixth and seventh panels show that males have friends who studied less in high school and expect to study less in college (compared to females), while blacks have friends who studied more in high school and expect to study more in college (compared to nonblacks).

The last panel of Table 3 describes friend study time. Consistent with own study time in Table 2, the first row shows that, on average, friend study time is 3.5 hours per day. The second and third rows of the last panel show that average friend study time is much lower for males than for females (3.16 vs. 3.76 hours per day).

Table 4 shows other network characteristics. Both the probability that a first-semester friendship no longer exists in the second semester and the probability that a second-semester

[8]Descriptive statistics about study time outcomes presented in Table 2 are computed at the level of individual study time reports, of which there may be up to four in each semester, for each student. When computing other descriptive statistics (including regressions), we use the semester-specific average (over the study time reports) for each student The two measures are very similar, other than the larger variance of the individual-report-based measure. As we make clear when we describe our estimation procedure, we use individual study time reports when estimating the structural parameters of our model.

[9]Pooling observations from both semesters, the first and last differences in means are significantly different at a 0.05 level and, given the relatively small number of black students, the middle difference in means is significant at a 0.10 level.

[10]Pooling observations from both semesters, all of these differences are significant at a 0.05 level.

[11]Therefore, the number of friends may exceed that elicited in the survey in Appendix A.1.

Table 1: Own summary statistics

Variable	Group	N	Mean	SD	Min	q1	q2	q3	Max
(1) Male indicator	all	307	0.44	0.5	0	0	0	1	1
	black	55	0.45	0.5	0	0	0	1	1
	nonblack	252	0.43	0.5	0	0	0	1	1
	above-med. HS GPA	155	0.33	0.47	0	0	0	1	1
	below-med. HS GPA	152	0.55	0.5	0	0	1	1	1
(2) Black indicator	all	307	0.18	0.38	0	0	0	0	1
	male	134	0.19	0.39	0	0	0	0	1
	female	173	0.17	0.38	0	0	0	0	1
	above-med. HS GPA	155	0.10	0.31	0	0	0	0	1
	below-med. HS GPA	152	0.26	0.44	0	0	0	1	1
(3) HS GPA	all	307	3.39	0.47	1.68	3.09	3.5	3.8	4
	male	134	3.24	0.51	1.68	2.9	3.21	3.7	4
	female	173	3.51	0.4	2.13	3.3	3.6	3.85	4
	black	55	3.14	0.46	2.24	2.78	3.1	3.52	4
	nonblack	252	3.45	0.46	1.68	3.19	3.53	3.8	4
	above-med. HS GPA	155	3.77	0.17	3.5	3.6	3.8	3.9	4
	below-med. HS GPA	152	3.00	0.35	1.68	2.8	3.08	3.29	3.47
(4) ACT	all	307	23.26	3.61	14	21	23	26	33
	male	134	22.54	3.77	14	20	23	25	31
	female	173	23.82	3.39	17	21	24	26	33
	black	55	19.91	2.51	14	18	20	21	25
	nonblack	252	23.99	3.4	14	22	24	26	33
	above-med. HS GPA	155	24.45	3.53	17	22	25	27	33
	below-med. HS GPA	152	22.04	3.28	14	20	22	24	31
(5) HS study	all	307	11.24	11.35	0	4	8	15	70
	male	134	11.43	11.94	0	3.12	8	15	70
	female	173	11.10	10.9	0	4	9	15	70
	black	55	15.29	14	0	5	10.5	20	70
	nonblack	252	10.36	10.51	0	3	7	14	70
	above-med. HS GPA	155	10.66	10.44	0	4	8	14.5	70
	below-med. HS GPA	152	11.84	12.21	0	3.38	8.25	15	70
(6) Expected study	all	307	24.96	11.61	0	17	23	31	64
	male	134	22.72	11.08	0.97	16	20.75	27.38	64
	female	173	26.68	11.74	0	19	25.5	33	57.5
	black	55	28.56	13.56	0	19	25	38.5	57.5
	nonblack	252	24.17	11.01	0	17	22.5	30.62	64
	above-med. HS GPA	155	25.18	10.47	0	18	23.5	32	56
	below-med. HS GPA	152	24.72	12.69	0	16	22.25	30.12	64

Note: The rows in each panel show descriptive statistics for the variable of interest in the first column, for the group in the second column. GPA is measured in GPA points (0-4). HS study and expected study are measured in hours/week.

friendship was not present in the first semester are 0.51. Consistent with the findings from Table 3, the correlations on the right side of the table show substantial sorting on the basis of observable characteristics.

Table 5 presents descriptive OLS regression results predicting own study time (left col-

Table 2: Own summary statistics for outcomes, by semester

Variable	Group	N	Mean	SD	Min	q1	q2	q3	Max
(1) Sem. 1 Own study	all	955	3.49	2.23	0	2	3.25	4.67	16
	male	401	3.23	2.38	0	1.67	3	4.33	14.67
	female	554	3.68	2.1	0	2	3.33	5	16
	black	158	3.83	2.23	0	2.33	3.41	5.33	11.67
	nonblack	797	3.43	2.23	0	2	3	4.67	16
	above-med. HS GPA	518	3.62	2.27	0	2	3.33	5	16
	below-med. HS GPA	437	3.34	2.17	0	2	3	4.67	14.67
(2) Sem. 2 Own study	all	945	3.5	2.12	0	2	3.33	4.67	14.33
	male	384	3.22	2.11	0	2	3	4.33	12
	female	561	3.7	2.11	0	2	3.33	5	14.33
	black	169	3.75	1.98	0	2.33	3.33	5	9.67
	nonblack	776	3.45	2.15	0	2	3.31	4.67	14.33
	above-med. HS GPA	513	3.66	2.06	0	2	3.33	5	12
	below-med. HS GPA	432	3.32	2.18	0	2	3	4.67	14.33
(3) Sem. 1 GPA	all	307	2.89	0.78	0	2.49	3.06	3.46	4.00
	male	134	2.72	0.80	0.30	2.17	2.80	3.29	4.00
	female	173	3.02	0.74	0	2.66	3.13	3.55	4.00
	black	55	2.42	0.78	0	1.82	2.57	2.84	4.00
	nonblack	252	3.00	0.74	0.3	2.58	3.11	3.55	4.00
	above-med. HS GPA	155	3.19	0.62	0.52	2.81	3.29	3.69	4.00
	below-med. HS GPA	152	2.59	0.8	0	2.00	2.66	3.12	4.00
(4) Sem. 2 GPA	all	301	2.93	0.78	0	2.53	3.05	3.46	4.00
	male	131	2.74	0.84	0	2.38	2.82	3.33	4.00
	female	170	3.07	0.71	0.44	2.66	3.20	3.54	4.00
	black	53	2.58	0.86	0.44	2.22	2.62	3.33	3.78
	nonblack	248	3	0.75	0.00	2.58	3.08	3.5	4.00
	above-med. HS GPA	155	3.21	0.66	0	2.82	3.36	3.74	4.00
	below-med. HS GPA	146	2.63	0.79	0.26	2.15	2.66	3.24	4.00

Note: The rows in each panel show descriptive statistics for the variable of interest in the first column, for the group in the second column. GPA is measured in GPA points (0-4). Own study is measured in hours/day and in this table is reported at the individual study report level.

umn) and GPA (right column), pooling observations over both semesters. The study time regression shows evidence of significant partial correlations of one's own study time (computed as the average amount the student reports studying in the time diaries within a semester) with own sex and own high school GPA. As for our study propensity measures, we estimate a positive, significant partial correlation between own study time and own high school study time. We do not estimate a significant correlation between own study time and expected study time when both propensity measures are included. However, when expected study time is the only study propensity measure included, we find that it has a positive, significant partial correlation with own study time (t-statistic of 2.2). The overall contribution of these two variables is substantial, with their omission reducing R-squared from 0.169 to 0.087 (see Table 13 in the appendix). Our novel measures of the propensity to study clearly have content. One's own study time also has a significant positive partial correlation with

Table 3: Average friend summary statistics, pooled over both semesters

Variable	Group	N	Mean	SD	Min	q1	q2	q3	Max
(1) Num. friends	all	614	3.31	1.58	1	2	3	4	10
	male	268	3.22	1.59	1	2	3	4	10
	female	346	3.38	1.57	1	2	3	4	9
	black	110	3.21	1.35	1	2	3	4	7
	nonblack	504	3.33	1.62	1	2	3	4	10
	above-med. HS GPA	310	3.34	1.62	1	2	3	4	10
	below-med. HS GPA	304	3.28	1.53	1	2	3	4	8
(2) Frac. male friends	all	614	0.43	0.39	0	0	0.33	0.75	1
	male	268	0.74	0.31	0	0.5	0.82	1	1
	not male	346	0.18	0.25	0	0	0	0.33	1
	black	110	0.43	0.4	0	0	0.33	0.83	1
	not black	504	0.42	0.39	0	0	0.33	0.75	1
	above-med. HS GPA	310	0.35	0.38	0	0	0.25	0.67	1
	below-med. HS GPA	304	0.5	0.39	0	0	0.5	1	1
(3) Frac. black friends	all	614	0.18	0.32	0	0	0	0.25	1
	male	268	0.18	0.32	0	0	0	0.25	1
	not male	346	0.17	0.33	0	0	0	0.2	1
	black	110	0.69	0.38	0	0.43	1	1	1
	not black	504	0.07	0.16	0	0	0	0	1
	above-med. HS GPA	310	0.10	0.22	0	0	0	0	1
	below-med. HS GPA	304	0.26	0.39	0	0	0	0.45	1
(4) Friend HS GPA	all	614	3.37	0.32	2.24	3.2	3.41	3.62	4
	male	268	3.29	0.33	2.25	3.07	3.34	3.53	4
	not male	346	3.44	0.29	2.24	3.29	3.46	3.64	4
	black	110	3.18	0.34	2.25	2.96	3.19	3.41	4
	not black	504	3.42	0.30	2.24	3.25	3.45	3.63	4
	above-med. HS GPA	310	3.46	0.27	2.65	3.29	3.46	3.63	4
	below-med. HS GPA	304	3.29	0.35	2.24	3.08	3.35	3.55	3.92
(5) Friend ACT	all	614	23.29	2.63	16	21.67	23.33	25	32
	male	268	22.72	2.64	16.33	21	23	24.64	31
	not male	346	23.74	2.54	16	22	23.67	25.5	32
	black	110	21.2	2.53	16	19.33	21	22.5	29
	not black	504	23.75	2.43	16.33	22.25	23.67	25.33	32
	above-med. HS GPA	310	23.79	2.42	17.5	22.23	23.67	25.33	32
	below-med. HS GPA	304	22.78	2.74	16	21	23	25	30
(6) Friend HS study	all	614	11.03	7.64	0	6	9.5	14.47	70
	male	268	10.53	7.37	0.5	5.17	9	14	37.33
	not male	346	11.41	7.83	0	6.5	9.79	14.6	70
	black	110	14.62	7.31	2.5	9.18	13.92	18.75	37
	not black	504	10.24	7.49	0	5.5	8.68	13.19	70
	above-med. HS GPA	310	11.48	8.44	0.5	6	9.7	14	70
	below-med. HS GPA	304	10.57	6.7	0	6	9.17	14.64	37.33
(7) Friend expected study	all	614	24.82	7.4	0	19.75	23.55	29.62	55
	male	268	22.89	6.97	4.06	18.23	21.65	27.05	55
	not male	346	26.33	7.38	0	21.02	25.06	31.38	52
	black	110	28.05	8.53	12	21.35	28.9	33.79	51
	not black	504	24.12	6.94	0	19.5	23	28.2	55
	above-med. HS GPA	310	24.72	7.42	0	20	23.55	29.48	55
	below-med. HS GPA	304	24.93	7.39	10.5	19.31	23.61	29.81	52
(8) Friend study	all	614	3.5	1.72	0	2.47	3.26	4.28	11.93
	male	268	3.16	1.49	0.5	2.21	3	3.88	8.46
	not male	346	3.76	1.83	0	2.65	3.51	4.5	11.93
	black	110	3.78	1.77	0.5	2.7	3.52	4.47	10.81
	not black	504	3.44	1.7	0	2.4	3.2	4.24	11.93
	above-med. HS GPA	310	3.64	1.79	0	2.56	3.36	4.41	11.93
	below-med. HS GPA	304	3.36	1.64	0.5	2.36	3.17	4.13	10.81

Note: The rows in each panel show descriptive statistics for the variable of interest in the first column, for the group in the second column. GPA is measured in GPA points (0-4). Own and friend HS study and expected study (top panel) are measured in hours/week. Friend study (bottom panel) is measured in hours/day. The variable "Friend z" for student i in period t is the average of the variable z across i's friends in period t.

friend study time (computed as the average over friends of their own study times). The GPA regression shows that own GPA has a significant positive partial correlation with being female, being nonblack, and having above-median high school GPA. Own GPA also has a significant partial correlation with own study time.

Table 4: Network characteristics

Friendship transitions	
Prob. friendship reported first semester but not second	0.51
Prob. second semester friendship is new	0.51

Note: Top row is computed according to $\Pr\{A_2(i,j) = 0|A_1(i,j) = 1\}$ and bottom row is computed according to $\Pr\{A_2(i,j) = 0|A_1(i,j) = 1\}$, where A_t is the adjacency matrix in semester t.

Correlations between own and avg. of friends	
Black	0.74
Male	0.71
HS GPA	0.23
Combined ACT	0.31
HS study time	0.23
Expected study time	0.14

Note: Each row is presents the correlation for a student's own measure and the average of their friends' measures, pooled over both semesters.

4 Model

Students are indexed by $i = 1, \ldots, N$ and time periods (semesters) by $t = 1, 2$. We denote the study time of student i in time period t as s_{it} and let S_t define a column vector collecting all students' study times during that period. We treat the adjacency matrix representing the network of friendships as pre-determined. This matrix in period t, denoted A_t, has a main diagonal of zeros and an (i, j) entry of one if student i has j as a friend and zero otherwise.[12] The average study time of i's friends during period t is

$$s_{-it} = \frac{\sum_{j=1}^{N} A_t(i,j) s_{jt}}{\sum_{j=1}^{N} A_t(i,j)}. \tag{1}$$

Taking into account their friends, students make decisions about how much to study in a particular semester by considering the costs and benefits of studying. The benefits of studying come from the accumulation of human capital. The production function for human capital, which we will also refer to as achievement, $y(\cdot)$, is:

$$y(s_{it}, \mu_{yi}) = \beta_1 + \beta_2 s_{it} + \mu_{yi}, \tag{2}$$

where μ_{yi} is a "human capital type" which allows the amount a person learns in school to vary across people, conditional on her own study level. As will be discussed in Section 5, in practice, this type will be constructed using observable characteristics that have consistently

[12]Other than its being full rank, we impose no restrictions on A_t. Though in our baseline empirical specification we use the union of reported links (i.e., $A_t(i, j) = 1$ if either i reports being friends with j, or vice versa), the model could also accommodate non-reciprocal links (i.e., i may link to j without j linking to i).

Table 5: Study time and GPA OLS regressions

	Dependent variable:	
	Own study	GPA
	(1)	(2)
Male	-0.369^{**}	-0.131^{*}
	(0.171)	(0.076)
Black	0.116	-0.225^{**}
	(0.214)	(0.109)
HS GPA	0.413^{**}	0.437^{***}
	(0.188)	(0.081)
ACT	-0.032	0.040^{***}
	(0.023)	(0.013)
HS study	0.043^{***}	0.001
	(0.008)	(0.004)
Expected study	-0.002	-0.006
	(0.009)	(0.003)
Friend study	0.166^{***}	
	(0.039)	
Own study		0.090^{***}
		(0.022)
Constant	1.915^{**}	0.417
	(0.759)	(0.362)
Observations	574	571
R^2	0.169	0.259

Note: *p<0.1; **p<0.05; ***p<0.01. Standard errors clustered at the student level are in parentheses. GPA is measured in GPA points (0-4). Own and friend HS study and expected study are measured in hours/week. Own and friend study are measured in hours/day. The variable "Friend z" for student i in period t is the average of the variable z across i's friends in period t.

been found to influence academic performance. We adopt a value-added formulation for the evolution of human capital, i.e., the human capital type is assumed to be a sufficient statistic for the history of prior inputs.

The cost of studying, $c(\cdot)$, is determined by:

$$c(s_{it}, s_{-it}, \mu_{si}) = \theta_1 s_{it} + \theta_2 \gamma(\mu_{si}) s_{it} + \frac{\theta_3 s_{it}}{s_{-it}^{\tau_s}} + \frac{\theta_4 \gamma(\mu_{si}) s_{it}}{s_{-it}^{\tau_s}} + \frac{\theta_5 s_{it}^2}{2 s_{-it}^{\tau_s}}, \tag{3}$$

where friend study time enters the cost function by reducing the cost of one's own studying, with curvature given by the exponent τ_s. As we show below, this cost function produces a best response function with desirable properties. Studying may be less arduous when one's friends are studying. We discuss below how this specification of the cost function is observationally equivalent with one in which social interactions are instead driven by conformity forces. The term μ_{si} is i's "study type," which allows the disutility from studying to vary across students, conditional on own and friend study levels. As will be discussed in Section 5, in practice, this type will be constructed from observable characteristics that help explains one's study time choices. Study types enter the model through $\gamma(\cdot)$. We define

$$\gamma(\mu_{si}) = \frac{1}{\exp(\tau_{\mu,1}\mu_{si} + \tau_{\mu,2}\mu_{si}^2)}, \tag{4}$$

which allows the cost function to have intercepts and slopes that vary across people of different study types. We refer to $\gamma(\mu_{si})$ as the "effective study type". We do not include a fixed cost of studying because very few students report zero study time over the semester.

With knowledge of $\{A_1, A_2\}$, all students' human capital types $\{\mu_{yi}\}_{i=1}^N$ and all students' study time types $\{\mu_{si}\}_{i=1}^N$, students simultaneously choose study times to maximize utility, which we assume to be separable across periods:[13]

$$u(s_{i1}, s_{i2}) = \left\{ \sum_{t=1}^{2} y(s_{it}, \mu_{yi}) - c(s_{it}, s_{-it}, \mu_{si}) \right\}. \tag{5}$$

Remark 1. *Before solving the model, it may be useful to include a brief discussion of what may seem to be the somewhat spare specification laid out thus far. When developing our model, we leaned on our intuition about what would be most important for generating social interactions in academic achievement in the first-year college context that we study, our proposed mechanism being that friends' study time choices affect one's own choice of study time and, thus, achievement. That being said, one strength of our unique data is that we are able*

[13] The alternative assumption, where students know only the current adjacency matrix when choosing their study times and calculate expectations over the future adjacency matrix, would have identical predictions in our model. See Section 4.1.

to test this specification against others that have received attention in the academic context, e.g., those including direct effects of peer characteristics in the production of achievement (contextual effects) and those including production complementarities. We discuss alternative specifications in Sections 4.2 and 7.1. As described in Section 7.1, our extensive testing of other specifications supports the parsimonious one we develop here.

4.1 Model Solution

Each student's decision problem is additively separable across time periods, meaning each student can solve each period's problem separately.[14] Student i's best response to friend study time in t is given by:

$$s_{it} = \underset{s \in [0,24]}{\arg\max} \{y(s, \mu_{yi}) - c(s, s_{-it}, \mu_{si})\}, \tag{6}$$

with the natural constraints that study time is nonnegative and cannot exceed 24 hours per day. The first order condition of (6) with respect to own study time yields $\frac{\partial y}{\partial s} = \frac{\partial c}{\partial s}$, i.e., the utility-maximizing study time equates the marginal return for increasing study time with the marginal cost. Expanding the first order condition gives:

$$\beta_2 = \theta_1 + \theta_2 \gamma(\mu_{si}) + \theta_3 \frac{1}{s_{-it}^{\tau_s}} + \theta_4 \frac{\gamma(\mu_{si})}{s_{-it}^{\tau_s}} + \theta_5 \frac{s_{it}}{s_{-it}^{\tau_s}}. \tag{7}$$

Solving for own study time yields the best response function, which expresses student i's study time as a function of friend study time, at an interior solution:

$$s_{it} = -\frac{\theta_3}{\theta_5} - \frac{\theta_4}{\theta_5}\gamma(\mu_{si}) + \frac{(\beta_2 - \theta_1)}{\theta_5} s_{-it}^{\tau_s} - \frac{\theta_2}{\theta_5}\gamma(\mu_{si}) s_{-it}^{\tau_s}. \tag{8}$$

Equation (3) shows that the term associated with θ_5 introduces curvature into the student's cost function. If θ_5 were zero, the student's objective in (6) would be linear in own study time and there would not exist an interior best response to friend study time. Equation (8) also shows that one of the preference parameters θ must be normalized. Therefore, we normalize θ_5 to one, which has the advantage of clearly showing that we allow for the possibility of finding no evidence of endogenous social interactions, which would occur if we estimated that both $(\beta_2 - \theta_1) = 0$ and $\theta_2 = 0$. The resulting final form of the student best response

[14] If utility were nonlinear in semester achievement or the argument of the cost function were study time over the whole year, the problem would no longer be separable across time periods. We assume student utility is linear in achievement because non-linearity of utility in achievement would be difficult to separate from non-linearity in the cost function without relying on functional form restrictions.

function for an interior solution is:

$$s_{it} = -\theta_3 - \theta_4\gamma(\mu_{si}) + (\beta_2 - \theta_1)s_{-it}^{\tau_s} - \theta_2\gamma(\mu_{si})s_{-it}^{\tau_s} \equiv \psi(s_{-it}, \mu_{si}). \tag{9}$$

Note that while best response functions depend on study type μ_{si}, it is sometimes notationally convenient to suppress the study type and write the best response function as $\psi_i(s_{-it})$. Own study time is increasing linearly in the productivity of own study time β_2 and may also be increasing in own effective study type, μ_{si}, depending on θ_2 and θ_4. We restrict parameters so that own study time has a strictly positive intercept and is a weakly increasing and weakly concave function of friend study time.[15]

As shown in Section 4.1.1, concave best response functions ensure existence of a unique equilibrium for the study time game. As shown in equations (8) and (9), the separable form we adopt for the cost function has the benefit of producing a closed-form solution for the student best response function. We show in Appendix B.1 that concavity of the best response function would result from any cost function possessing the natural properties of being strictly convex in s_{it} and weakly concave in s_{-it}.

4.1.1 Equilibrium

Definition 1 (Period Nash equilibrium). *A pure strategy Nash equilibrium in study times $S^* = [s_1^*, s_2^*, \cdots, s_N^*]'$ satisfies $s_i^* = \psi(s_{-i}^*, \mu_{si})$, for $i \in N$, given adjacency matrix A.*

Claim 1. *Let k be a number strictly greater than 24. There exists a unique pure strategy Nash equilibrium if $\psi_i : R^N \mapsto R$ are weakly concave and weakly increasing, $\psi_i(0) > 0$, and $\psi_i(k) < k$ for $i \in N$.*

Proof. See Appendix B.2. □

We compute the equilibrium by iterating best responses.

4.2 Model Discussion

4.2.1 Friend Study Time

We define friend study time as the average study times of one's friends. Our framework could also accommodate specifications where friend study time was defined to be the total study

[15] The strictly positive intercept restriction corresponds to $\min_{i\in N}\{-\theta_3 - \theta_4\gamma(\mu_{si})\} > 0$. The weakly increasing restriction corresponds to $\min_{i\in N}\{(\beta_2 - \theta_1) - \theta_2\gamma(\mu_{si})\} \geq 0$. Weak concavity corresponds to further requiring $\tau_s \leq 1$. These restrictions, combined with $s_{it} \leq 24$, are sufficient to have the well-behaved equilibrium described in Section 4.1.1. In practice, however, we are able to estimate the model using weaker restrictions, described in Section 5.

time of one's friends. Indeed, we previously estimated a specification of our model in which the denominator of (1) was $\alpha \sum_{j=1}^{N} A_t(i,j) + (1-\alpha)$, where $\alpha \in [0,1]$, $\alpha = 1$ corresponds to the average of friends' study times, and $\alpha = 0$ corresponds to the total of friends' study times. We define friend study time to be the average of friends' study times because we found α to be 1.

4.2.2 Other Mechanisms Generating Endogenous Social Interactions

Conformity Our specification of the cost function allows friend study effort to reduce one's own cost of studying. Others have allowed social interactions to emerge from a cost of deviating from peer actions, i.e., from a force producing conformity (see, e.g., Brock and Durlauf (2001), Moffitt (2001), Blume et al. (2015)). We show in Appendix B.3.1 that such a specification would be observationally equivalent to the one we adopt.

Production Complementarities Another proposed mechanism is that social interactions arise through production complementarities, where increases in peer inputs increase the marginal product of one's own input (e.g., Calvó-Armengol et al. (2009)). From a conceptual standpoint, the decision to specify our model without production complementarities was informed by the notion that friends in the first year of college may spend relatively little time talking about coursework, with some empirical support for this provided by Stinebrickner and Stinebrickner (2006).[16]

That being said, we show in Appendix B.3.2 that in the typical case, where one only had data on either the input (e.g., study effort) or output (e.g., achievement), our cost-reduction-based specification (or, equivalently, a conformity-based specification) would be observationally equivalent to a specification exhibiting production complementarities. This point is also made by Blume et al. (2015). Therefore, because we measure both inputs and outcomes, we are in the unique position to examine the potential roles played by cost reductions and production complementarities. As we discuss in Section 7.4, we do not find evidence for such a mechanism in our context.

4.2.3 Dynamic Behavior

We assume the human capital type is constant between the periods. Though it is feasible to extend our static framework to a dynamic framework allowing the human capital type to evolve between periods, the benefits of doing this are mitigated by two facts: (1) each

[16]Stinebrickner and Stinebrickner (2006) find that students spend very little time talking about coursework with their roommates; it is not a big leap to imagine the same would be true of students and their friends.

model period corresponds to a semester, which is shorter than the period typically considered when estimating value-added production functions in an educational context (see Hanushek (1979) and Todd and Wolpin (2003) for discussions of issues related to the estimation of education production functions), and (2) we study students during their freshman year, which, under the liberal-arts curriculum at Berea, is typically before they start taking specialized course material (meaning second semester coursework does not build heavily on first semester coursework). Consistent with these facts, as we discuss in Section 7.5, we find that out-of-sample outcomes, simulated from parameters estimated on only first-semester data, fit second-semester data quite well.

5 Estimation

The model provides a mapping from the adjacency matrix A_t and all the students' types $\{(\mu_{si}, \mu_{yi})\}_{i=1}^N$ to a unique equilibrium in study times for all students, S_t^*. The equilibrium study times S_t^* generate achievement in equilibrium y_{it}^*, via the production function $y(s_i, \mu_{yi})$. The model is operationalized by parameterizing a student's types as linear combinations of observable characteristics collected in a vector x_i. That is, $\mu_{si} = x_i'\omega_s$ and $\mu_{yi} = x_i'\omega_y$, where the parameter vectors ω_s and ω_y respectively determine study and human capital types.[17] The vector x_i includes indicators for being black and being male, along with high school GPA, combined ACT score, average hours per week of study time in high school, and expected hours per week of study time in college. This allows us to express each student's equilibrium study time and achievement as a function of A_t and all students' characteristics, which we collect in a matrix X. Given the full set of (model) parameters $\Gamma = (\beta_1, \beta_2, \theta_1, \theta_2, \theta_3, \theta_4, \omega_s, \omega_y, \tau_{\mu,1}, \tau_{\mu,2}, \tau_s)'$, we write these outcomes for individual i as

$$s_{it}^* = \psi(s_{-it}^*, \mu_{si}) = \delta_i^s(A_t, X; \Gamma) \tag{10}$$

and

$$y_{it}^* = y(s_{it}^*, \mu_{yi}) = \delta_i^y(A_t, X; \Gamma), \tag{11}$$

where s_{-it}^* is defined by applying equation (1) to S_t^* and A_t.

Our measure of achievement, denoted $\widetilde{y}_{it}$, is the student's semester grade point average (GPA), which is measured on a four-point scale. In our data, 7% of student-semester observations have a GPA of four and 1% have a GPA of zero. Therefore, we take a Tobit approach to modeling GPA. We define latent GPA as $y_{it}^* + \eta_{yit}$, where η_{yit} is a Gaussian measurement error that is IID and independent from A and X. Our Tobit model, with censoring at zero

[17] We set coefficient on high school GPA in the study type $\omega_{s,\text{HS GPA}} = 1$ to identify $\gamma(\cdot)$.

and four, is

$$\widetilde{y}_{it} = \begin{cases} 4 & \text{if } y^*_{it} + \eta_{yit} \geq 4 \\ 0 & \text{if } y^*_{it} + \eta_{yit} \leq 0 \\ y^*_{it} + \eta_{yit} & \text{otherwise.} \end{cases} \tag{12}$$

The GPA component of the likelihood function for individual i at time t is the likelihood for this Tobit model:

$$L^y_{it} = \Phi\left(\frac{0 - \delta^y_i(A_t, X; \Gamma)}{\sigma_{\eta_y}}\right)^{\mathbf{1}\{\widetilde{y}_{it}=0\}} \times \left(1 - \Phi\left(\frac{4 - \delta^y_i(A_t, X; \Gamma)}{\sigma_{\eta_y}}\right)\right)^{\mathbf{1}\{\widetilde{y}_{it}=4\}} \times \frac{1}{\sigma_{\eta_y}}\phi\left(\frac{\widetilde{y}_{it} - \delta^y_i(A_t, X; \Gamma)}{\sigma_{\eta_y}}\right), \tag{13}$$

where Φ and ϕ denote the CDF and PDF, respectively, of the standard normal distribution.

The likelihood function also takes into account study time outcomes. Our measures of s^*_{it} come from up to four 24-hour time diaries completed by each student i in semester t. We use R_{it} to denote the set of reports for student i in semester t. Study time report r for student i in semester t is denoted $\tilde{s}_{rit}$, and is allowed to be a noisy measure of s^*_{it}.[18] Because approximately 5% of our study time observations are zero, we use a Tobit approach for reported study time. Defining latent study time as $s^*_{it} + \eta_{srit}$, reported study time is

$$\widetilde{s}_{rit} = \begin{cases} 0 & \text{if } s^*_{it} + \eta_{srit} \leq 0 \\ s^*_{it} + \eta_{srit} & \text{otherwise.} \end{cases} \tag{14}$$

The likelihood contribution for report r of student i in semester t is

$$L^s_{rit} = \Phi\left(\frac{0 - \delta^s_i(A_t, X; \Gamma)}{\sigma_{\eta_s}}\right)^{\mathbf{1}\{\widetilde{s}_{rit}=0\}} \times \frac{1}{\sigma_{\eta_s}}\phi\left(\frac{\widetilde{s}_{rit} - \delta^s_i(A_t, X; \Gamma)}{\sigma_{\eta_s}}\right). \tag{15}$$

The total likelihood contribution for student i is therefore[19]

$$L_i = \left(\prod_t \prod_{r \in R_{it}} L^s_{rit}\right) \times \left(\prod_t L^y_{it}\right), \tag{16}$$

the sum of which across students we maximize to obtain our estimated parameters.[20]

[18] Stinebrickner and Stinebrickner (2004) document how reported study time varies within semesters.

[19] We allow for dependence within student when computing standard errors.

[20] Note that the theoretical model assumes best response functions are strictly positive, nondecreasing, and weakly concave. These restrictions are difficult to directly impose in terms of restrictions on the parameter space when there is heterogeneity in best response functions. Therefore, we adopt an indirect approach, of verifying whether best response functions derived from posited parameters satisfy the restrictions. Specifically, when estimating the model, we use the weaker restrictions that the 75th percentile effective study type's best-response function is nonnegative and that equilibrium study times are strictly positive; we also impose the natural upper bound on daily study time (24 hours a day). As we show in Section 7, none of these restrictions are close to binding at our estimated parameters.

6 Specification Test

This section develops a specification test that has power against alternative data generating processes that have unobserved determinants of study time. It presents the test statistic, shows how to calculate the residuals used to compute it, and shows how to decompose these residuals in a way that facilitates our analysis of when the test has power. We develop the test assuming there is one study time report and one period, which allows us to drop the corresponding subscripts, and that there is no censoring in observed study times. We show how we implement our test using more than one study time report and more than one period in Section 7.2.

Let $\widehat{\Gamma}$ denote the vector of estimated parameters and let $\tilde{s}_i$ denote i's reported study time. The study time residual for student i is

$$\widehat{\eta}_{si} = \widetilde{s}_i - \delta_i^s(A, X; \widehat{\Gamma}), \tag{17}$$

and the average of i's friends' residuals is

$$\widehat{\eta}_{s,-i} = \frac{\sum_{j=1}^{N} A(i,j)\widehat{\eta}_{sj}}{\sum_{j=1}^{N} A(i,j)}. \tag{18}$$

Consider the following regression of a student's own residual on the average of her friends' residuals:

$$\widehat{\eta}_{si} = a + b\widehat{\eta}_{s,-i} + \xi. \tag{19}$$

Under the null hypothesis of proper specification, the error terms in our study time regression are IID, so the true value of b is zero. Our test statistic is simply the t-statistic for a test of $b = 0$ in regression (19), $\widehat{b}/SE(\widehat{b})$, which has a limiting standard normal distribution under the null.

Our claim is that this specification test will have power against alternatives where omitted variables are present. Intuitively, if there are important unobserved variables influencing students' decisions, they will generally induce cross-sectional correlation across students because they will enter students' best responses in equilibrium. Therefore, an absence of correlation in residuals across students is consistent with a lack of omitted variables. We show this more formally now.

Consider the following scenario with an (potentially) omitted variable. We examine the special case with $\tau_s = 1$. As we show in Section 7, this case is consistent with our baseline empirical results, where we find that best response functions are linear.

Using a composite parameter $c \equiv [\beta_2 - \theta_1]$, define the subset of parameters identified by just equation (9), the student's policy function, as $\Gamma_2 = (c, \theta_2, \theta_3, \theta_4, \omega_s, \tau_{\mu,1}, \tau_{\mu,2})'$. To simplify notation, we refer to the terms in (9), $[-\theta_3 - \theta_4\gamma(\mu_{si})]$ and $[(c - \theta_2\gamma(\mu_{si}))\, s_{-i}]$, as $f_1(x_i; \Gamma_2)$ and $f_2(x_i; \Gamma_2)$, respectively. As in Section 5, x_i contains student i's observed characteristics, which enter the policy function through effective study type $\gamma(\mu_{si})$. The equation for an individual student is

$$s_i = f_1(x_i; \Gamma_2) + f_2(x_i; \Gamma_2)s_{-i}. \tag{20}$$

In order to represent the system of equations for all students in a vector S, use $F_1(X; \Gamma_2)$ to denote a column vector stacking the $f_1(x_i; \Gamma_2)$ for all i. We use the notation $W(X; \Gamma_2)$ for a matrix that has zeros in the same positions as the zeros in A and nonzero entries in locations where A has ones. In place of the ones in row i of A, $W(X; \Gamma_2)$ contains

$$\frac{1}{\sum_{j=1}^{N} A(i,j)}[f_2(x_i; \Gamma_2)]. \tag{21}$$

The system of equations is thus:

$$S = F_1(X; \Gamma_2) + W(X; \Gamma_2)S. \tag{22}$$

Note that (22) is simply a re-written version of the baseline model we developed in Section 4. Solving for S, we obtain the equilibrium vector of study times

$$S^* = (I - W(X; \Gamma_2))^{-1}[F_1(X; \Gamma_2)], \tag{23}$$

where the right side corresponds to the vector stacking $\delta_i^s(A, X; \Gamma)$ for all students.[21]

Incorporating our measurement error η_s, we obtain the data generating process for observed study time $\widetilde{S}$ under the *null hypothesis* of correct specification:

$$\widetilde{S} = (I - W(X; \Gamma_2))^{-1}[F_1(X; \Gamma_2)] + \eta_s. \tag{24}$$

Now consider an alternative in which the model was misspecified. In particular, suppose a vector of characteristics V was omitted by the econometrician but was observed by all students, entering the best response system in the following manner:

$$S = F_1(X; \Gamma_2) + W(X; \Gamma_2)S + V. \tag{25}$$

Again solving for S, the equilibrium system of equations has the following form:

$$S^* = (I - W(X; \Gamma_2))^{-1}[F_1(X; \Gamma_2) + V]. \tag{26}$$

[21] Note that, as it expresses outcomes as a reduced form in terms of X and A, the specification test would not be affected by the inclusion of "contextual effects" entering in the typical additively separable manner.

In general, the matrix $(I - W(X;\Gamma_2))^{-1}$ will have many non-zero entries because students will typically be directly or indirectly connected to many other students. Therefore, many, if not all, elements of V will influence a given student's equilibrium study time in this alternative.

Decompose V into two components according to:

$$V = \Pi(X) + u, \tag{27}$$

where we assume that u is mean independent of X (and, hence, $F_1(X;\Gamma_2)$) and $W(X;\Gamma_2)$. We are agnostic about correlation patterns in u across students; in particular, friends may have correlated u. Consider, for example, a scenario where male and female students have the same expected value of u, but where a (mean-zero) sex-specific shock induces correlations among students of the same sex, who are likely to be friends with each other. Substituting this expression for V into (26) and incorporating our measurement error η_s gives the data generating process for observed study time under the *alternative hypothesis*:

$$\begin{aligned}\widetilde{S} =& (I - W(X;\Gamma_2))^{-1}[F_1(X;\Gamma_2) + \Pi(X) + u] + \eta_s \\ =& (I - W(X;\Gamma_2))^{-1}[F_1(X;\Gamma_2) + \Pi(X)] + (I - W(X;\Gamma_2))^{-1}u + \eta_s. \end{aligned} \tag{28}$$

It is convenient to re-write (28) with a composite error ϵ:

$$\begin{aligned}\widetilde{S} =& (I - W(X;\Gamma_2))^{-1}[F_1(X;\Gamma_2)] + \epsilon, \qquad (29)\\ \text{where } \epsilon =& (I - W(X;\Gamma_2))^{-1}\Pi(X) + (I - W(X;\Gamma_2))^{-1}u + \eta_s. \end{aligned}$$

By definition, residuals must be computed using estimates of Γ_2, rather than the true value. To derive the residuals $\widetilde{\epsilon}$, consider the least squares estimator of Γ_2 in the study time regression, $\widehat{\Gamma}_2$, i.e., the estimate of Γ_2 that minimizes $\epsilon'\epsilon$ in (29). The fitted values for $\widetilde{S}$ using $\widehat{\Gamma}_2$ are $(I - W(X;\widehat{\Gamma}_2))^{-1}[F_1(X;\widehat{\Gamma}_2)]$. Let $\widetilde{\Gamma}_2$ denote the probability limit of $\widehat{\Gamma}_2$. In large samples, the fitted values of $\widetilde{S}$ based on our estimator would then be $(I - W(X;\widetilde{\Gamma}_2))^{-1}[F_1(X;\widetilde{\Gamma}_2)]$, which we can add and subtract from (29), resulting in

$$\widetilde{S} = (I - W(X;\widetilde{\Gamma}_2))^{-1}[F_1(X;\widetilde{\Gamma}_2)] + \widetilde{\epsilon}, \tag{30}$$

where

$$\begin{aligned}\widetilde{\epsilon} =& \underbrace{\{(I - W(X;\Gamma_2))^{-1}[F_1(X;\Gamma_2) + \Pi(X)] - (I - W(X;\widetilde{\Gamma}_2))^{-1}[F_1(X;\widetilde{\Gamma}_2)]\}}_{\text{“prediction bias”}} \\ &+ \underbrace{(I - W(X;\Gamma_2))^{-1}u}_{\text{equilibrium propagation of } u} + \eta_s. \end{aligned} \tag{31}$$

The first term ("prediction bias") in $\widetilde{\epsilon}$ is due to the omission of $\Pi(\cdot)$, i.e., it represents a mean misspecification in (30). As we show below, there will not necessarily be a prediction bias. The second term is due to the influence of u upon equilibrium study effort. This second term is what our specification test is designed to detect. $\widehat{\Gamma}_2$ will, in general, be inconsistent for Γ_2 if $\Pi(\cdot) \neq 0$.

In general, our test will have power, i.e., the ability to detect the type of alternative (25), because the error $\widetilde{\epsilon}$ will exhibit cross-sectional correlation when $V \neq 0$. We show this by considering cases (i) without prediction bias and (ii) with prediction bias.

There would be no prediction bias (i.e., we would be in case (i)) if there exists a $\ddot{\Gamma}_2$ such that $F_1(X; \ddot{\Gamma}_2)$ nests $F_1(X; \Gamma_2) + \Pi(X)$.[22] In principle, this nesting could be accomplished by adopting a sufficiently flexible functional form for $F_1(\cdot; \cdot)$. Therefore, in this case, although elements of Γ_2 may be inconsistently estimated (i.e., plim $\widehat{\Gamma}_2 \neq \Gamma_2$), residuals obtained from (30) would only include components based on u and η_s. That is, bias in $\widehat{\Gamma}_2$ would not pervade to the residuals.[23] Of course, in practice, it may be necessary to impose restrictions on $F_1(X; \cdot)$, meaning there may potentially be prediction bias in study time. We discuss this in case (ii).

Case (i): No prediction bias: Consider first the case with no prediction bias, leaving us to focus on the u component of $\widetilde{\epsilon}$ in (31). In general, the term $(I - W(X; \Gamma_2))^{-1}u$ will exhibit cross-sectional dependence because its elements are linear combinations of many of the components of u.

In order for there to be no cross-sectional covariance in $(I - W(X; \Gamma_2))^{-1}u$, the shocks u would need to have a covariance matrix that was orthogonalized by $(I - W(X; \Gamma_2))^{-1}$. For example, consider the case where u was generated according to

$$u = (I - W(X; \Gamma_2))e, \tag{32}$$

with e IID and $\mathrm{E}[ee'] = I$. For reasonable ranges of $W(X; \Gamma_2)$ in our application, such a u process would possess strong negative correlations among closely linked students. For example, consider our point estimate for Γ_2, which we present in Section 7, and our baseline adjacency matrix for the first semester, A_1. For the process in (32), in order for u to be orthogonalized by $(I - W(X; \Gamma_2))^{-1}$, the ratio of the average covariance of u between friends to the average variance across students would have to be -0.31.

[22] This is because the u are mean independent from X and $W(X; \Gamma_2)$, which means that the least-squares estimate of Γ_2, which minimizes $\epsilon'\epsilon$, would minimize the prediction bias component of $\widetilde{\epsilon}$.

[23] For example, this would be true in the commonly considered case where we can write $F_1(X; \Delta_1) = X\Delta_1$ and $\Pi(X) = X\Delta_2$, where Δ_1 and Δ_2 are matrices of parameters, in which case the estimated $\widehat{\Delta}_1$ would have plim $\widehat{\Delta}_1 = \Delta_1 + \Delta_2$.

The main focus of the literature is on the case of positive assortative matching.[24] Therefore, we believe such a negative correlation is not the most salient one. Further note that the necessary orthogonalization could not occur when u are independent in the cross section. Moreover, even in the presence of negative cross-sectional correlations in u, only specific correlation structures could produce the necessary orthogonalization. In our example above, negative covariances in u that were either larger or smaller than -0.31 would generate correlated residuals ($\widetilde{\epsilon}$). In summary, our test will have power to detect the omitted variables u as long as they do not have very specific covariance structures.

Case (ii): Prediction bias: In the case where there is a prediction bias in study time (which can only occur if there is an omitted variable bias, i.e., $\Pi(\cdot)$ is nonzero), our test would not have power if the first term in equation (31), the prediction bias term, offset cross-sectional correlations in $(I-W(X;\Gamma_2))^{-1}u$, such that there would be no cross-sectional covariance in ϵ. For example, negative covariances in the bias term could, in principle, exactly cancel with the positive covariances that we anticipate in $(I - W(X;\Gamma_2))^{-1}u$. Our strong prior is that this scenario is implausible, due to the positive covariances across friends in their values of x_i and the bias term being a smooth function of x_i. Intuitively, because friends have similar observed characteristics (x_i), the prediction errors of students and their friends—which would only exist due to the inability of $F_1(X;\cdot)$ to fit study time for students with certain observed characteristics—will likely be positively correlated. Most importantly, *prediction bias would have to exactly cancel out the u component* to not have power against the alternative hypothesis. Such a problematic scenario would be a knife-edge case.

To make the test developed above more concrete, in Appendix E we develop an example environment with dyadic, separate networks with homogeneous best responses. In addition to simplifying notation, the specification developed in Appendix E corresponds to case (i), i.e., there is no prediction bias.

7 Estimation Results

Table 6 contains parameter estimates. The top panel presents the parameters that enter the achievement production function. The key parameter is the marginal product of own study time on achievement, β_2. The point estimate of 0.254 implies that increasing own study time by one hour per day increases achievement by about a quarter of a GPA point, ceteris

[24] Epple and Romano (2011) contains a thorough discussion of sorting in the presence of peer effects. Zeitlin (2011) studies peer effects in a social learning context, finding that own and friend information shocks are negatively correlated. This finding is unsurprising in a learning environment, where one may gain more when one's friends have different information.

paribus. It is reassuring that this result is quantitatively similar to that from Stinebrickner and Stinebrickner (2008a), who estimate that, for freshman at Berea, an extra hour per day of studying would increase GPA by 0.36 points (with a standard error of 0.183 points), using whether a randomly assigned roommate brought a video game as a shifter for one's own study time. Table 6 shows that students with high GPAs in high school and high ACT scores have significantly higher human capital, and black students have significantly lower human capital.

Table 6: Parameter Estimates

Parameter	Estimate	SE	Description
Production function*			
β_1	-0.350	0.4185	intercept
β_2	0.254	0.0651	marginal product of own study time
$\omega_{y,\text{HS GPA}}$	0.470	0.0808	coefficient on HS GPA in human capital type
$\omega_{y,\text{ACT}}$	0.047	0.0112	coefficient on ACT in human capital type
$\omega_{y,\text{Black}}$	-0.213	0.1074	coefficient on Black in human capital type
$\omega_{y,\text{Male}}$	-0.037	0.0849	coefficient on Male in human capital type
$\omega_{y,\text{HS study}}$	-0.007	0.0042	coefficient on HS study in human capital type
$\omega_{y,\text{expected study}}$	-0.005	0.0035	coefficient on expected study in human capital type
Study cost function / Best response function, setting $\tau_s = 1$**			
θ_1	-1.074	0.1551	affects common best response slope
θ_2	0.874	0.2351	affects heterogeneity in best response slope
θ_3	-0.907	0.8097	affects common best response intercept
θ_4	0.096	1.2800	affects heterogeneity in best response intercept
$\tau_{\mu,1}$	0.105	0.0601	linear term for study type
$\tau_{\mu,2}$	-0.003	0.0028	quadratic term for study type
$\omega_{s,\text{HS GPA}}$	1.000	–	coefficient on HS GPA in study type, fixed to 1
$\omega_{s,\text{ACT}}$	-0.063	0.0870	coefficient on ACT in study type
$\omega_{s,\text{Black}}$	-0.735	0.7459	coefficient on Black in study type
$\omega_{s,\text{Male}}$	-1.065	0.7892	coefficient on Male in study type
$\omega_{s,\text{HS study}}$	0.344	0.1554	coefficient on HS study in study type
$\omega_{s,\text{expected study}}$	0.005	0.0309	coefficient on expected study in study type
Shocks			
σ_{η_y}	0.721	0.0185	sd measurement error for human capital
σ_{η_s}	2.159	0.0377	sd measurement error for observed study time

* Production function: $y = \beta_1 + \beta_2 s_{it} + \mu_{yi}$, where $\mu_{yi} = x_i' \omega_y$.

** Best response function: $s_{it} = -\theta_3 - \theta_4 \gamma(\mu_{si}) + (\beta_2 - \theta_1) s_{-it} - \theta_2 \gamma(\mu_{si}) s_{-it}$, where $\mu_{si} = x_i' \omega_s$. Recall that we allowed for $\tau_s \in [0, 1]$ in our estimation, but, finding it to be indistinguishable from 1, we fixed $\tau_s = 1$ and re-estimated.

As can be seen in equation (9), the curvature in the best response function is given by τ_s, the exponent on s_{-it}. We estimated the model allowing τ_s to be in the set [0,1], nesting the assumption of a linear best response function (i.e., that τ_s=1).[25] However, because our initial estimation provided evidence that τ_s is indistinguishable from 1, we re-estimated the model fixing τ_s=1.

Estimates of the parameters in the study cost function appear in the second panel of Table 6. To ease their interpretation, we substitute them into the best response function, yielding

$$\widehat{\psi}(s_{-it}, \widehat{\mu}_{si}) = \{0.907 - 0.096\widehat{\gamma}(\widehat{\mu}_{si})\} + \{1.328 - 0.874\widehat{\gamma}(\widehat{\mu}_{si})\}\, s_{-it}. \tag{33}$$

The first bracketed term in equation (33) represents the intercept of the best response function for student i, i.e., how much this student would study even if her friends did not study at all. This term consists of $-\theta_3$ =0.907, the common component of the intercept across students, and $-0.096\widehat{\gamma}(\widehat{\mu_{si}})$, the component characterizing variation in the intercept across students. Likewise, the second bracketed term in equation (33) reveals the slope, or *reactiveness*, of the best response function, that is, how a student's choice of study time depends on the study time of her friends. This term consists of $(\beta_2 - \theta_1) = 1.328$, the common component of the slope across students (we estimate the composite parameter $(\beta_2 - \theta_1)$ to have a standard error of 0.1804), and $-0.874\widehat{\gamma}(\widehat{\mu}_{si})$, the component characterizing variation in the slope across students. The negative point estimate for θ_1 means that the common slope component of student study times was higher than could be explained by only the marginal product of study time in producing achievement (β_2). With $\gamma(\mu_{si}) = \frac{1}{\exp(\tau_{\mu,1}\mu_{si}+\tau_{\mu,2}\mu_{si}^2)}$, the latter component in both the first and second bracketed terms depends on the estimated values of $\tau_{\mu,1} = 0.105$ and $\tau_{\mu,2} = -0.003$, which indicate that $\gamma(\cdot)$ is decreasing and convex in one's study type, μ_s. In turn, the value of one's study type, μ_s, is determined by the cost function parameters ω_s. As seen at the end of Table 6, study type is increasing in high school GPA and high school study time, but is smaller for males.[26]

To provide a better sense of the total effect of peer study effort in the best response functions, Figure 1 plots best response functions for several effective study types, $\widehat{\gamma}(\widehat{\mu}_s)$: the lowest (lower dotted green line), 25th percentile (lower dashed purple line), median (dot-dashed red line), and 75th percentile (higher dashed purple line), and the highest (higher dotted green line). The table just below Figure 1 calculates equation (33) for each of these effective study types, presenting the type-specific intercept (i.e., the first bracketed term

[25] See Blume et al. (2015) for an extensive discussion of linear social interactions models. We have verified that (the row-normalized) A_t are not idempotent, facilitating estimation of linear best response functions.

[26] Though black students study considerably more than nonblack students, the coefficient on being black is negative. Black students have much higher high school study levels, which we find to be an important determinant of study type.

in equation (33)) in the top row and the coefficient on friend study time (i.e., the second bracketed term in equation (33)) in the bottom row. The first row shows that there is little heterogeneity in the intercepts of best response functions. The second row shows that reactiveness to peer study time is increasing in effective study type. The effect of peer study time is significantly positive for all effective study types, including the lowest effective study type (lower dotted green line, or first column of the table).[27] Combining the slope and intercept terms, one's optimal study choice is increasing in study type. That is, the best response is always increasing in s_{-it} and is often substantial.[28]

To get a sense of whether the estimated heterogeneity in reactiveness is significant, in Table 7 we present 95% confidence intervals for differences in best response slopes for different groups of students. Females have significantly steeper best response functions than males, students with above-median high school GPAs have significantly steeper best response functions than those with below-median high school GPAs, and students with above-median high school study time have significantly steeper best response functions than those with below-median high school study time.

Table 7: Means and 95% confidence intervals for difference in slope of best response function, by group

Comparison	Mean	2.5%	97.5%
Female-Male	0.058	0.015	0.115
Black-Nonblack	0.014	-0.013	0.044
High HS GPA-Low HS GPA	0.034	0.015	0.056
High Study HS-Low Study HS	0.177	0.077	0.299

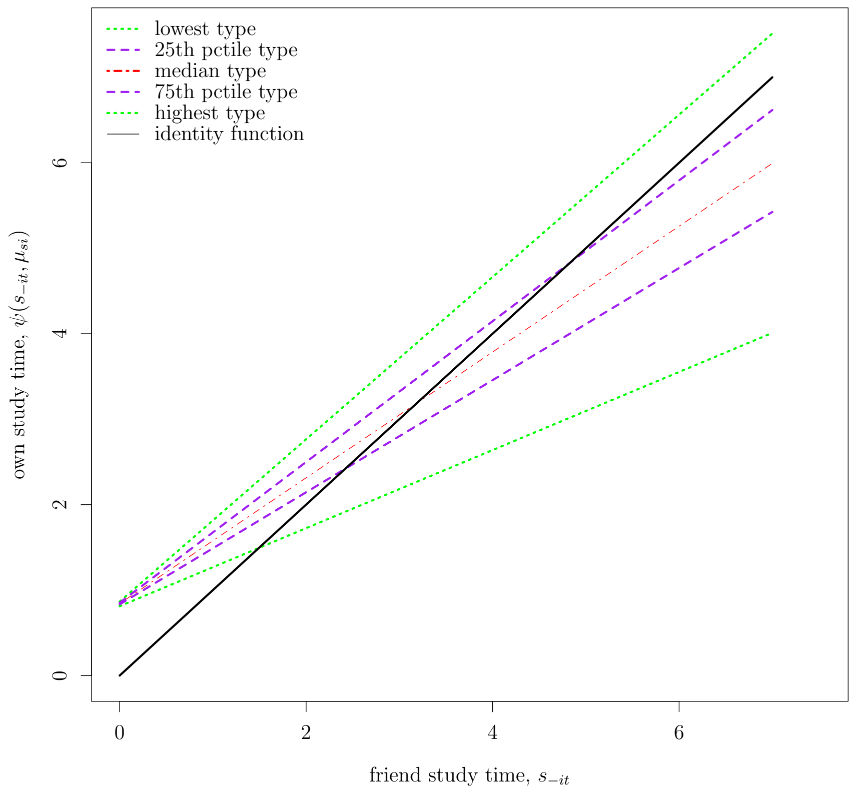
"High-" and "Low HS GPA" respectively refer to above- and below-median high school GPA. "High-" and "Low Study HS" respectively refer to above- and below-median high school study time.

In addition to describing individual heterogeneity in best response functions, Figure 1 provides evidence about the implications of this heterogeneity. To see this, note that the intersection of each best response function with the identity function indicates the equilibrium study outcome in a hypothetical scenario in which a student and someone of the same effective study type were paired. Therefore, by comparing where the different types' best response functions intersect the identity function (solid black line), we can identify equilibria

[27] The flexible specification we have developed to allow for heterogeneity in best response functions makes it difficult to discern whether the slopes of best response functions are significantly positive via direct examination of parameters in Table 6. Therefore, we adopted a conservative approach to assess statistical significance. We computed the 95% confidence interval for the best response slope for each student and then examined whether any of these confidence intervals contained zero; the lower bound on the union of these confidence intervals is 0.11.

[28] As noted in Section 5, we did not need to impose that best response functions are increasing in estimation.

Figure 1: Estimated study best response functions for different effective study types $\widehat{\gamma}(\widehat{\mu}_s)$



Effective study type $\widehat{\gamma}(\widehat{\mu}_s)$:	Lowest	25th pctile	Median	75th pctile	Highest
Intercept	0.81	0.83	0.84	0.85	0.87
Coefficient on s_{-it}	0.46	0.66	0.74	0.82	0.95

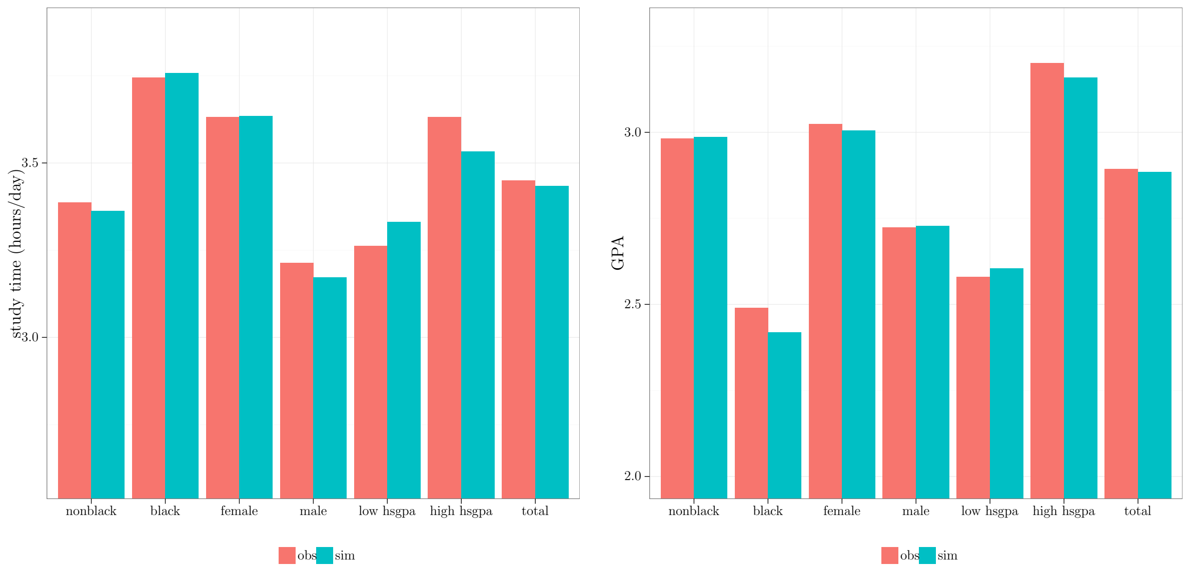
Note: Own and friend study times are measured in hours/day. Each column represents the estimated best response function for an effective study type. For example, the middle column indicates that the median effective study type has the estimated best response function $s_{it} = 0.84 + 0.74s_{-it}$.

when each student is matched with someone of her respective study type. When two 75th percentile effective study types are paired they would study almost 5 hours each, almost twice the amount two 25th percentile types would study when paired. Our estimates indicate that the game exhibits a complementarity. If matched by study type, students may study more in total, and therefore, have higher total achievement. However, whether students will take advantage of this complementarity depends on how they sort into friendships.

Figure 2 shows that the model closely fits mean observed study time (left panel) and GPA (right panel), both in total and by student characteristics.[29] Even though the relationship

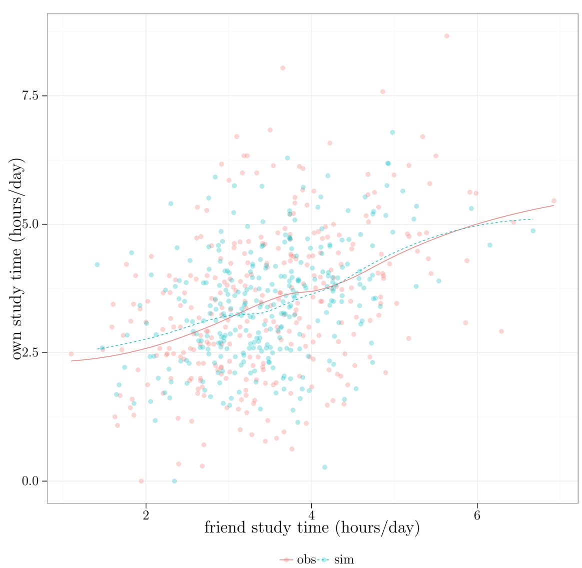
[29]Model outcomes are simulated by first solving for equilibrium outcomes given $\widehat{\Gamma}$ and then applying measurement errors, using the specification in Section 5.

Figure 2: Fit of mean study time (left) and GPA (right), by group



Note: "obs" are means computed using the data and "sim" are means of outcomes simulated from the model. "High-" and "Low HS GPA" respectively refer to above- and below-median high school GPA.

Figure 3: Fit of own study time against friend study time



Note: "obs" correspond to data and "sim" correspond to model simulations. Each point corresponds to a pair of own and friend study time (both are measured in hours/day). The lines are fitted values from a local quadratic regression. For each value of friend study time the fit is computed using the closest 75% of the observations via weighted least squares, with weights proportional to $(1-(\text{distance}/\text{max. distance})^3)^3$). See stat_smooth in the R package ggplot2 for details (Wickham (2009), R Core Team (2015)).

between own and friend study time is not explicitly targeted (i.e., friend study time outcomes do not enter the likelihood), the model also closely captures this relationship. Figure 3 plots own versus friend study time, for both the data (solid red line) and simulated outcomes (dashed blue line).

In the remainder of this section we discuss potential endogeneity problems, present the results from our specification test, and present evidence about the robustness of our estimates.

7.1 Endogeneity

Our primary endogeneity concerns arise from the potential for the relationship between a student's study effort and that of her peers to be due, in part, to friendships being formed on the basis of potentially unobserved determinants of study time. One possible concern is that the relationship between own and friend study time is driven by institutional factors. One prominent example is that if students in science courses tend to study more and befriend students in their courses, there may be a spurious relationship between own and friend study time. We find that a version of the descriptive regression in Table 5, including both own and friend fraction of courses which are science, does not appreciably change the partial correlation between own and friend study time (0.166 vs. 0.160).[30] This is not surprising given that students may make friends outside their classes, the large majority of curriculum choices for freshman are required general or introductory classes, and there is not much variation in the number of classes taken.[31] In the same vein, dormitories are not specialized at Berea (e.g., there are not "study" dormitories or separate dormitories for student athletes).

Perhaps a more important concern is that students arrive at school with differing propensities to study, which affects how they sort into friendships. We address this concern by taking advantage of our survey collection to obtain direct measures of students' propensities to study. Our baseline survey elicited information about: 1) how much a student expected to study in college and 2) how much a student studied in high school. As we discussed in Section 3.2, these measures of the propensity to study clearly have content, as they are strongly correlated with how much a student studies. We stress a crucial feature of this information on study propensity is that our survey design allowed this information to be collected immediately after students arrived on campus, before students could be influenced by their friendships at Berea.

As is always the case, it is difficult to know *a priori* whether observable characteristics

[30] See Table 14 in the appendix.

[31] On average, students take about one additional course in their area of specialization per semester in their freshman year.

can address potential endogeneity concerns. Therefore, we next present results from our specification test, which was designed to detect a wide variety of unobserved determinants of study time, in particular, those underlying endogeneity concerns.

7.2 Specification Test Results

To simplify exposition, in Section 6 we developed our specification test for one period and one study time report. This section starts by showing how we implement the test using our data for two periods (semesters) and multiple study time reports. Recall that predicted equilibrium study time for student i in semester t is $\delta_i^s(A_t, X; \widehat{\Gamma})$.[32] We define i's semester-t study time residual as the average residual over i's semester-t study time reports, $\widetilde{s}_{rit}$:

$$\widehat{\eta}_{sit} \equiv \frac{1}{\sum_{r \in R_{it}} 1} \sum_{r \in R_{it}} \left(\widetilde{s}_{rit} - \delta_i^s(A_t, X; \widehat{\Gamma}) \right) = \overline{\widetilde{s}}_{it} - \delta_i^s(A_t, X; \widehat{\Gamma}), \tag{34}$$

where $\overline{\widetilde{s}}_{it}$ is i's average study time over reports r in semester t. To implement the test, we average student's residuals over both semesters, i.e., $\widehat{\eta}_{si} = \frac{\widehat{\eta}_{si1} + \widehat{\eta}_{si2}}{2}$. We then compute the average of friends' average residuals for each student in each semester according to $\widehat{\eta}_{s,-it} = \frac{\sum_{j=1}^{N} A_t(i,j) \widehat{\eta}_{sj}}{\sum_{j=1}^{N} A_t(i,j)}$.

Our test statistic is the t-statistic for a test of zero slope in a regression of $\widehat{\eta}_{si}$ on $\widehat{\eta}_{s,-it}$, pooled across semesters. Under the baseline specification, in which our new measures of study propensity (high school study time and expected study time) enter students' study types, our test statistic has a p-value of 0.716, corresponding to a correlation between own and friend study time residuals of 0.017.[33] Thus, our test results suggest that our model is well-specified. There is no evidence of an endogeneity problem arising from students positively sorting into friendships based on unobserved determinants of study time.

In the presence of an omitted characteristic that generates an endogeneity problem, our test should indicate a positive relationship in own and friends' residuals. To demonstrate that our test can detect a relationship in such a scenario, we construct an example where there is likely an endogeneity problem, by estimating a restricted version of our model in which we purposefully omit our novel measures of study propensity. Notably, this restricted specification uses only measures of student characteristics that are typically available to researchers. Because our empirical results show that these measures are both determinants of study time and also related to our measures of incoming human capital and friendship

[32]To be consistent with our specification test, we also re-estimate $\widehat{\Gamma}$ assuming there is no censoring in study time and then use this estimate to compute the test statistic. Given the very small number of observations that were censored, the estimated $\widehat{\Gamma}$ (and corresponding test statistics) are virtually identical between the censored and non-censored versions. Details are available upon request.

[33]We use a distribution approximation for this test that allows correlation within students.

choices, their omission should generate a correlation across friends' residuals. The estimated correlation in this scenario is 0.208, and the test statistic for a slope coefficient of zero has a p-value of 1.12e-6, providing strong evidence against the null hypothesis of a slope of zero.

Taking these two residual correlations together, our test results suggest that our new measures of study propensity play a crucial role in addressing endogeneity concerns in our context.

Note that failing to reject the null model does not imply that we perfectly predict observed student study effort. Rather, failing to reject the null hypothesis means there are not significant determinants of study time that would introduce detectable correlations between own and friend residuals. For example, independently distributed reporting errors, e.g., of the type we assume in Section 5, would create a divergence between model and reported study time.

7.3 Human Capital Spillovers ("Contextual Effects")

We have focused on a mechanism wherein friend study time may affect one's own study time, which in turn may affect one's achievement via a production function. An alternative mechanism often considered in the literature involves peer characteristics directly entering the achievement production technology ("contextual effects"). For example, friends with high human capital may provide quick and reliable answers to questions, or may know more about specific course requirements, generating human capital spillovers.

Our *a priori* belief that a model without such spillovers may be quite natural is directly related to the mechanism for social interactions that we examine. In the short run that we study, it seems reasonable to believe that the primary reason a student's academic performance would be related to a particular observable characteristic of her friends is that the student's time-use is influenced by the good (or bad) study habits of friends with these characteristics. Models estimated without study time information would label this relationship as "contextual effects". In contrast, in our approach, which is made possible by the collection of time-use information, this relationship would be explicitly accounted for by our proposed mechanism in which one's study time is influenced by the study time of one's peers, thus removing a channel that would otherwise be labeled as "contextual effects".

Nevertheless, because, in theory, there could be spillovers not captured by our proposed mechanism, it is prudent to thoroughly examine whether friend characteristics explain achievement, even after accounting for our mechanism of interest. As a starting point, we augment the GPA prediction regression in Table 5, which regressed own GPA on own characteristics and own study time, by adding friend characteristics. Consistent with the results shown in Table 5, own study time remains a significant predictor of own GPA. However, friend characteristics are not significant predictors of students' own GPAs; adding friend characteristics increases the adjusted R^2 of the regression from 0.246 to 0.247.[34]

Given the literature's interest in the direct transmission of peer characteristics and our ability to separately identify them from endogenous social interactions, it is worthwhile to look beyond the prediction equation described above, and estimate specifications allowing for friend characteristics to directly affect one's achievement, i.e., human capital spillovers in the production function. To this end, we re-estimated the model using two alternative specifications. First, since a human capital spillover would naturally emerge from friend human capital types, we extend the technology (2) to allow for direct achievement transmission via human capital types:

$$y(s_{it}, \mu_{yi}) = \beta_1 + \beta_2 s_{it} + \beta_{3,\text{cont}} \mu_{y,-i,t} + \mu_{yi}, \tag{35}$$

where $\mu_{y,-i,t} \equiv \frac{\sum_{j=1}^{N} A_t(i,j)\mu_{yj}}{\sum_{j=1}^{N} A_t(i,j)}$, i.e., the average of period-t-friend human capital types. This specification's parsimoniousness makes it attractive from a practical level, but it is also conceptually attractive, as one would naturally expect friends with higher-than-predicted achievement (i.e., those with higher own human capital types, μ_{yj}) to be those who would also transmit more achievement to their friends. In this specification, a human capital spillover in the production of student achievement would correspond to $\beta_{3,\text{cont}} \neq 0$. As discussed in detail in Appendix C.1, we fail to reject that $\beta_{3,\text{cont}}$ is zero, with a point estimate of $\widehat{\beta}_{3,\text{cont}} = 0.111$ that has an accompanying standard error of 0.138.

Second, to examine the importance of the restriction implied by (35), that the determinants of one's own human capital type and one's human capital spillovers on one's friends are the same (to scale), we also re-estimated the model allowing for a more flexible specification based on a "contextual human capital type". Specifically, we define student i's contextual human capital type according to $\mu_{y,\text{cont},i} = x_i' \omega_{y,\text{cont}}$, where $\omega_{y,\text{cont}}$ is a vector containing six new parameters (one for each characteristic entering human capital and study types), and extend the technology (2) to be

$$y(s_{it}, \mu_{yi}) = \beta_1 + \beta_2 s_{it} + \mu_{yi} + \mu_{y,\text{cont},-i,t}, \tag{36}$$

where $\mu_{y,\text{cont},-i,t} \equiv \frac{\sum_{j=1}^{N} A_t(i,j)\mu_{y,\text{cont},j}}{\sum_{j=1}^{N} A_t(i,j)}$, i.e., the average of period-t-friend contextual human capital types. In this more flexible specification, a human capital spillover in the production of student achievement would correspond to one of the parameters in $\omega_{y,\text{cont}} \neq 0$. As dis-

[34]For example, the partial correlation coefficients for friend high school GPA, friend combined ACT score, and friend high school study have t-statistics of 0.75, .286, and -0.5, respectively.

cussed in Appendix C.1, we fail to reject that the vector $\omega_{y,\text{cont}}$ is zero at any conventional significance level (the likelihood ratio test statistic has a p-value of 0.368). The element of $\omega_{y,\text{cont}}$ with the most explanatory power is the parameter related to the share of male friends, which has a point estimate of $\widehat{\omega}_{y,\text{cont},Male} = 0.189$, though its standard error of 0.123 renders it statistically indistinguishable from zero.

Because we have data on both inputs and outcomes, there is more than one place in which contextual effects could enter our model. We chose the above extensions instead of, say, including a direct effect of friend characteristics in the cost function (3), because the results from our specification test do not provide strong evidence of omitted characteristics in the determination of study time choices.[35] These results lead us to conclude that mechanisms involving a direct role of friends characteristics in explaining achievement are not motivated in our application.[36] Given our *a priori* belief that this would be the case, we have thus chosen to retain the specification without human capital spillovers as our baseline specification.

7.4 Production Complementarities

As we show in Appendix B.3.2, our data on study time inputs and achievement outcomes allow us to separately identify production complementarities from cost-based mechanisms (which, as we show in Appendix B.3.1, are observationally equivalent to conformity-based mechanisms). As we discussed in Section 4.2.2, prior research suggests that production complementarities may not be very large, because students are not obliged to talk about coursework with their friends. However, given our unique ability to separately identify them from other proposed mechanisms, it is prudent to examine the potential role they play in determining study time and achievement. Our starting point is to add an interaction of own and friend study time to the GPA prediction regression in Table 5. We find that the estimated interaction between own and friend study time is insignificant and very small, with a partial correlation coefficient of 0.005 and standard error of 0.01.

Going beyond the above prediction regression, we also re-estimated the model using a specification that extends the technology (2) to be

$$y(s_{it}, \mu_{yi}, s_{-it}) = \beta_1 + \beta_2 s_{it} + \beta_{3,\text{comp}} \frac{s_{it}}{s_{-it}} + \mu_{yi}, \tag{37}$$

[35] Moreover, from a conceptual standpoint, it is *a priori* not obvious why friend characteristics, such as their high school GPA, would relate to one's own study time choices, after taking into account how much friends study.

[36] This is consistent with other work in which authors have noted that it may not be reasonable to expect large effects in the short run, given that students may not be taking the same classes and given that students often to not spend a large amount of time studying together.

where s_{it} is own study effort and s_{-it} is friend study effort. If $\beta_{3,comp.} < 0$, then increases in peer effort increase the marginal product of one's own effort.

As we discuss in Appendix C.2, we fail to reject that $\beta_{3,\text{comp}}$ is zero, with a point estimate of $\widehat{\beta}_{3,comp} = 0.904$ that has an accompanying standard error of 0.633. This estimation result led us to retain the specification without production complementarities in the technology as our baseline.

7.5 Dynamic Behavior and Model Validation

We discussed in Section 4.2.3 why it seemed reasonable to assume that the human capital type μ_{yi} was constant across semesters. In principle, however, first-semester achievement could increase students' human capital coming into the second semester, in which case, a model estimated using first-semester data may have difficulty fitting second-semester outcomes. On the other hand, if the out-of-sample fit turned out to be good, this would suggest a limit to the potential improvement in model fit from the addition of first-semester human capital or other dynamic considerations. Indeed, such an exercise could also be useful in discerning, more generally, whether the assumed micro-structure of our model (e.g., functional form assumptions, etc.) does a reasonably good job of capturing the key moving parts in our context.

Based on the above reasoning, we conducted an out-of-sample validation exercise by re-estimating our model using only data from the first semester and seeing how well it fit second-semester outcomes. As discussed in Appendix C.3, the second-semester fit is good. Therefore, we conclude that the static model may be appropriate for the relatively short time frame we consider in this paper.

8 Quantitative Findings

How much does it matter who your friends are? We use our estimated model to conduct two counterfactual exercises. First, we characterize how students respond to changes in friend study time by exogenously increasing (shocking) the study time of each student and measuring how outcomes would change for other students in the network. In addition to providing evidence about how network structure and student characteristics jointly determine how students are affected by their peers, this exercise provides a natural framework for quantifying the importance of equilibrium effects as well as the importance of heterogeneity in the effect of peers. Second, because peer effects are a function of not only how students respond to changes in peer inputs but also who is friends with whom, we examine how outcomes would differ if, instead of sorting into friendships as shown in Table 4, students were randomly assigned friends. This exercise provides a natural comparison point from which we can assess the importance of homophily in friendships.

Throughout this section, we compare outcomes between baseline and counterfactual scenarios for achievement, own study time, and friend study time. We use s_{it}^{cf} and s_{it}^{baseline} to denote student i's study time in the counterfactual and baseline scenarios, respectively. We define the treatment effect on achievement for student i in period t as $\Delta_{it}^{y} \equiv y(s_{it}^{\text{cf}}, \mu_{yi}) - y(s_{it}^{\text{baseline}}, \mu_{yi})$. Treatment effects for own and friend study time are defined analogously.

8.1 Network Structure, Student Characteristics, and the Response to Peer Input Changes

To provide quantitative evidence about how students respond to changes in peer study time, we estimate the impulse response to an impulse of increasing study effort. Specifically, we increase (shock) the study time of a single student by one hour per day in a particular semester and examine the responses of all other students in the network in that semester. We summarize our findings when we perform this exercise 614 times (once for each of the 307 students in each of the two semesters).

The averages in the first row of Table 8 show how the mean effect of the study shock evaluated at the new equilibrium, i.e., taking into account the full set of feedback effects in the network, varies with a student's distance from the shocked student. For example, to obtain the number in the second column we first compute, for each student j in each of the two semesters t, the mean response in achievement for all students who are one link away from j when j is shocked in semester t. Averaging this mean response over all shocked students j and both semesters shows that students who are one link away from the shocked student have an average achievement gain of 0.078 GPA points. Similarly, the third, fourth, and fifth columns, respectively, show that students who are two links, three links, and four links away from the shocked student, respectively, have average achievement gains of 0.022, 0.006, and 0.002 GPA points, respectively. The final column involves first computing, for each student j in each of the two semesters t, the *total* response in achievement, $\sum_{i \neq j} \Delta_{it}^{y}$, for all students (other than j) who are in the network when j is shocked in semester t. Averaging this total response over all students and semesters shows that, on average, the total effect of the shock is 0.52 GPA points.

Effects evaluated at the new equilibrium will be larger than partial equilibrium effects, which only take into account how the shock to a student influences students who are directly linked to her (i.e., iterating best response functions once). To quantify the importance of this difference, the second row of Table 8 shows the partial equilibrium effects. The average effect on students who are one link away from the shocked student is about 1/4 smaller under partial equilibrium than when than under the new equilibrium (0.059 vs. 0.078 GPA points), while, by definition, the effect on the (typically) large number of students who are two or more links away from the shocked student is zero in the partial equilibrium case. The last column shows that, on average, the total response of the shock is only 0.19 GPA points. Therefore, if we considered only partial equilibrium effects we would, on average, understate the achievement response by 64%.

Table 8: Average change in achievement (GPA points)

	Avg. response, by distance from shocked node					Total
Dist. from shocked stud.:	0	1	2	3	4	response
New equilibrium	0.254	0.078	0.022	0.006	0.002	0.52
Partial equilibrium	0.254	0.059	0.000	0.000	0.000	0.19

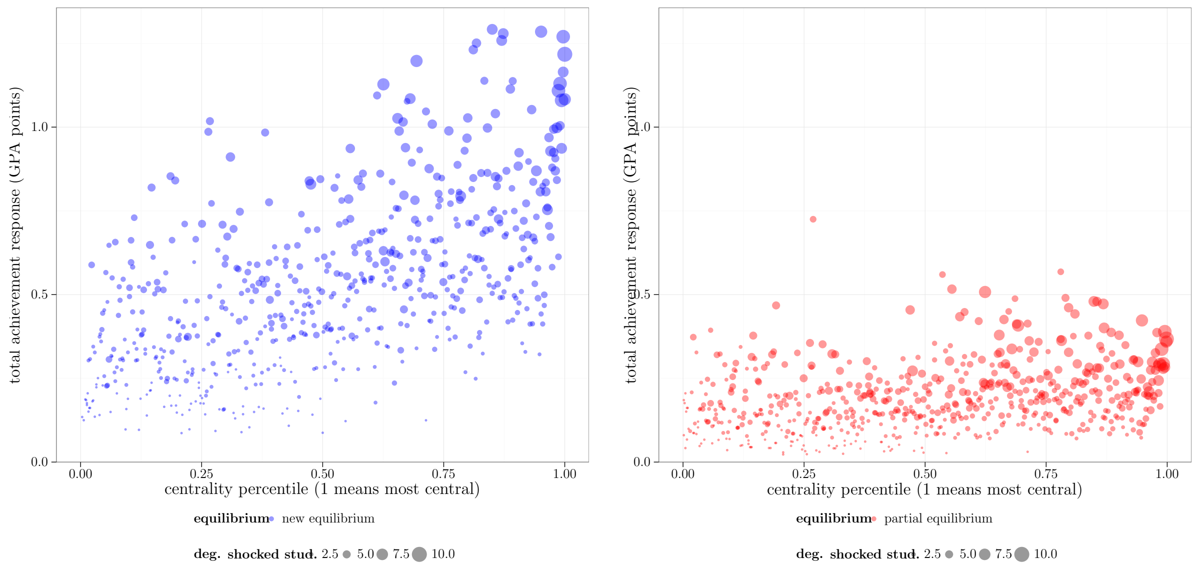
Note: The top row presents the mean effect on achievement (averaging over shocked students and semesters) at the new equilibrium, by distance from shocked student, where the shocked student has distance 0. The bottom row presents the mean effect on GPA immediately due to the impulse, by distance from shocked student. The mean total response, in the last column, is the average GPA response to shocking students j over periods t, excluding the effect on the shocked student, i.e., $\frac{1}{\sum_{j,t} 1} \sum_{j,t} \left(\sum_{i \neq j} \Delta_{it}^{y} \right)$.

We next examine how much the total response $\sum_{i \neq j} \Delta_{it}^{y}$ varies, depending on which student j is shocked in t. We find that the total response in achievement varies substantially depending on which student is shocked. For example, the first quartile, median, and third quartile of the total increase in achievement at the new equilibrium are 0.33, 0.49, and 0.66 GPA points, respectively. To get a better sense of why shocking different students can produce such different gains, the left panel of Figure 4 shows the relationship between the centrality of the shocked student and the total response at the new equilibrium.[37] As before, this calculation excludes the mechanical gain in achievement experienced by the shocked student. Each dot records the total achievement response (y-axis) by the percentile centrality that semester, i.e., by how central the shocked student is (x-axis). The size of each (blue) dot shows the degree (i.e., number of friends) of the shocked student. Larger dots are concentrated at the top-right, and smaller ones at the bottom-left. That is, students with more friends tend to have higher centrality indices and larger achievement gains. Intuitively, because the effects of effort changes are stronger the closer students are, the total response is higher when the shocked student is more centrally located.[38] The right panel of Figure 4 plots

[37]We use what is called a "closeness" centrality measure, given by the reciprocal of the sum of shortest distances between that student and every other student in the graph. Average distance to others for unconnected students is set to the number of students (Csardi and Nepusz (2006), Freeman (1979)).

[38]The notion that certain students may disproportionately affect other students is related to the concept

Figure 4: Total achievement response (GPA points), by centrality of shocked student



Note: The vertical location of each dot represents the total achievement response to shocking a different student; the left panel presents the total gain at the new equilibrium and the right panel presents the partial equilibrium total gain. The x-axis indicates the shocked student's centrality to other students and dot size denotes the degree of the shocked student.

partial equilibrium effects (red dots). We can see here that, though shocked students have the same degree (dot sizes), the average response is not as strongly increasing in centrality of the shocked student. This is the case because the equilibrium effects play a larger role the more densely connected the shocked student is to the rest of the network.

Figure 4 evinces variation in the total achievement response (i.e., the y-axis) to shocking different students who are similarly central (i.e., the x-axis) and who also have the same number of friends (i.e., dot sizes). We use two examples to illustrate how the structure of the social network interacts with the distribution of best response functions to determine how changes in students' actions affect other students.

The left panel of Figure 5a shows the subgraph containing students within three degrees of the student whose shock creates the largest total achievement response (1.29 GPA points). The right panel shows the subgraph containing students within three degrees of the student whose shock creates the smallest total achievement response (0.087 GPA points). In each case, the shocked student is denoted by a red star. Squares represent males and circles represent females. Shapes corresponding to black students are shaded and those corresponding to nonblacks are unshaded. The area of the circle or square representing a student other than the shocked student is proportional to the slope of that student's best response

of a "key player", studied in Ballester et al. (2006).

function, where larger shapes correspond to more reactive students. Both subgraphs show homophilous sorting: black students tend to be friends with other black students (and nonblacks with nonblacks), males tend to be friends with males (and females with females). In general, students with steeper best response functions tend to be friends with each other.

Differences in the total response can be due to differences in link structure and how heterogeneous students are arranged on the network. The link structures of the subgraphs are very different. The shocked student in the left panel has more friends (6 vs. 1) and more students within three degrees (39 vs. 12).[39] In addition to the structure of links, how the heterogeneous students are arranged on the network matters. Although the average slope of best response functions is roughly similar between the subgraphs, 0.759 in the left vs. 0.698 in the right, the friends of the shocked student in the left panel have steeper best response functions than the friend of the shocked student in the right panel. In the right panel, the shock is immediately dampened by being passed through the student's only, relatively nonreactive friend.

Figure 5b shows the analogous plot, where the area of the shape is now proportional to the achievement gain for that student. The effect of the shock dies off in the same pattern illustrated by the first row of Table 8, that is, shapes further from the star tend to be smaller. Friends of the shocked student in the left subgraph gain much more than the friend of the shocked student in the right subgraph. Due to the much steeper best response functions of the shocked student's friends, the impulse dies out much less quickly in the left subgraph. Indeed, the gains for students who are two links from the shocked student in the left subgraph are about as large as the gain for the student directly connected to the shocked student in the right subgraph. This persistence comes from both the steeper best response functions of direct friends of the shocked student and the fact that many of them are also connected to each other, further augmenting the effects of the shock through feedback. This implies the effectiveness of policies targeting students may depend critically on how they fit into the arrangement of the social network.[40]

8.2 The Effect of Sorting into Friendships

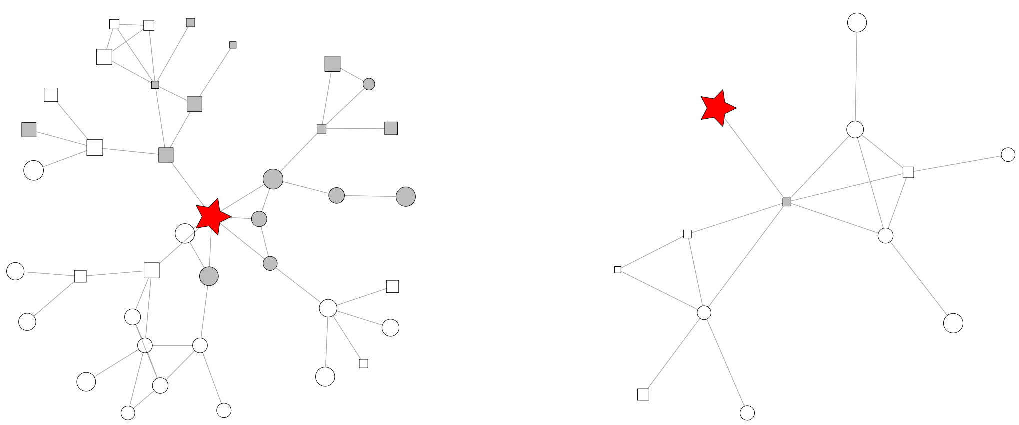
Section 8.1 studied how students respond to the input choices of others, taking into account the baseline network, which exhibits homophily. To directly examine homophily and, therefore, provide further evidence about the importance of peers, we compare achievement under the baseline social network with achievement under a counterfactual where friends are homo-

[39]We limit this illustration to students within three degrees based on the first row of Table 8, which shows the total impact dies off quite quickly in distance from the shocked student.

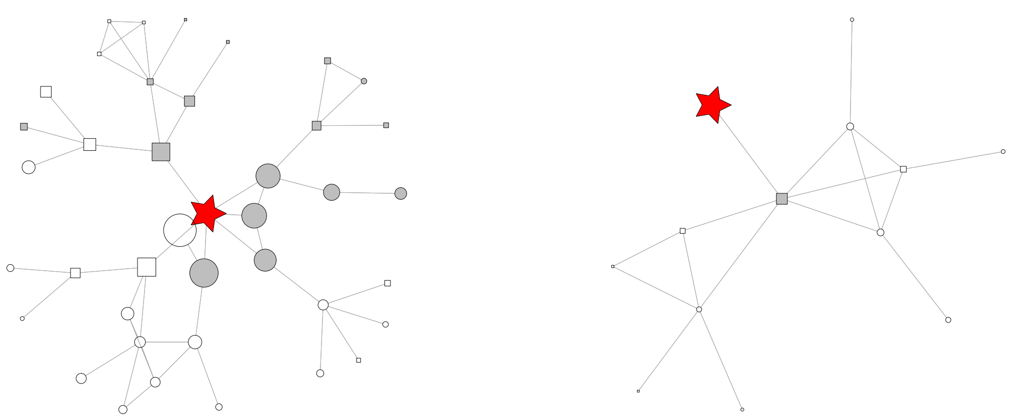
[40]See Fryer Jr (2011) for an example in which students are incentivized based on inputs to achievement.

Figure 5: Subgraphs corresponding to students producing the largest and smallest total achievement responses

(a) Slope of best response functions for students within three degrees of the student producing largest total response when shocked (left) and smallest total response when shocked (right)



(b) Gain in achievement for students within three degrees of the student producing largest total response when shocked (left) and smallest total response when shocked (right)



Note: Red star indicates shocked student, males are square (females are circles), blacks are shaded (nonblacks are unshaded), and area of squares and circles is proportional to outcome of interest for corresponding students (i.e., (a) slope of best response function or (b) gain in achievement from shocking starred student) geneously distributed across students. In this counterfactual, for each semester, we maintain the marginal distribution of friends per student observed in the data, but replace reported links with random draws from the entire sample of students. We then form a counterfactual symmetrized A matrix in the same manner as it was formed for the actual data, as described in Section 3. Repeating this process 300 times for each of the two semesters produces 300 pairs of simulated adjacency matrices.[41]

Table 9 summarizes changes in model outcomes between the baseline and counterfactual, averaged over all 300 simulated networks. Achievement is measured in GPA points and study times are in hours per day. The first column shows the average change in study time, across all students and all simulated networks, that results from moving to homogeneous (i.e., randomly assigned) friends. The first row shows that, on average, moving to this counterfactual would reduce own study time by 0.10 hours. Intuitively, students who in reality (i.e., under the baseline) have friends with high study types are most harmed by the move to a homogeneous distribution, which makes them much more likely to have lower study type friends. This explains why females, blacks, and students with above-median high school GPAs, who tend to be high study types and are seen in Table 4 to often have friends with high-study-type characteristics under the baseline, see own study time fall by 0.20, 0.25, and 0.15 hours, respectively. Conversely, males, who have less studious peers under the baseline, tend to study more when friends are homogenized. Importantly, the estimated complementarities, which arise due to the heterogeneity in best response functions combined with sorting into friendships based on effective study type, imply that the gains of lower study types are smaller than the losses of the higher study types. This explains the overall decrease in own study time. Removing the sorting in the manner of our experiment does not merely re-allocate output, but also lowers total output. Accordingly, the standard deviation of own study time drops by 30%. A similar story drives both the overall results and the stratified results associated with changes in friend study time in the second column of Table 9.

The third column of Table 9 shows the average change in achievement across all students and all simulated networks that result from the changes in study time found in the first column. The first row shows that, on average, moving to the counterfactual would reduce achievement by 0.02 GPA points. However, as expected given the findings of study time, the declines are largest for black students, female students, and students with above-median high

[41] For example, in the first semester the algorithm starts with IID draws of counterfactual "friends per student" from the empirical marginal distribution of friends per student in A_1, divided by two and rounded to the nearest integer, because A_1 has been union-symmetrized. The number of directed links per student is set to the student's "friends per student" draw. Directed links are IID draws from the whole set of other students.

school GPAs. As before, the losses to these groups are not offset by the gains to other groups. Homogenizing the distribution of friends' characteristics would increase the baseline GPA gap between nonblack and black students of 0.5 GPA points by 14%, reduce the baseline GPA gap between female and male students of 0.31 GPA points by almost 20%, and reduce the baseline GPA gap between students with above-median and below-median high school GPAs of 0.60 GPA points by 7%. Overall, homogenizing friends would reduce the standard deviation of achievement by 5%.

To gauge whether the effects reported above are significantly different from the baseline, we also report the range of the change in mean achievement, by group discussed above, across simulations. Figure 6 illustrates the 2.5th and 97.5th percentiles of the mean change for each group, along with the average mean changes presented in Col. (3) of Table 9. Intervals for black students, females, and students with above-median high school GPAs are all well below zero, as is the interval for the total change in achievement; only the group with below-median high school GPAs has an interval that contains zero. The significant effects indicated by these intervals reinforce our finding that sorting significantly affects student achievement.[42]

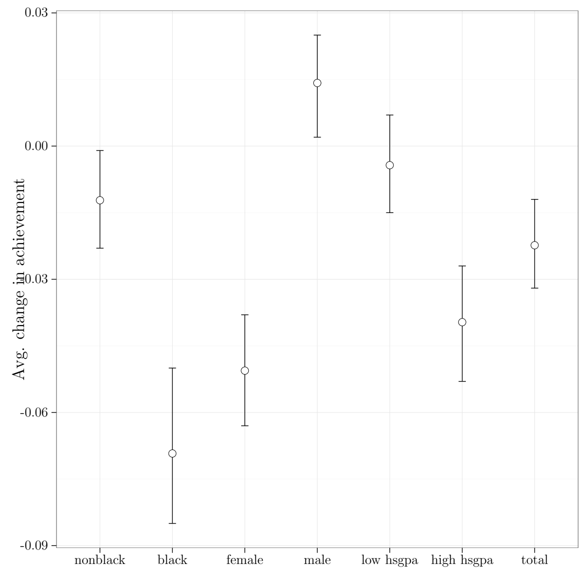
Table 9: Average changes for study time (hours/day) and achievement (GPA points) resulting from counterfactual homogeneous distribution of friend characteristics, across simulated networks

	Own study time (1)	Friend study time (2)	Achievement (3)
Total	-0.10	-0.09	-0.02
Nonblack	-0.07	-0.04	-0.01
Black	-0.25	-0.36	-0.07
Female	-0.20	-0.25	-0.05
Male	0.05	0.12	0.01
Below-med. HS GPA	-0.03	0.02	0.00
Above-med. HS GPA	-0.15	-0.20	-0.04

Note: Means are computed over simulated networks.

[42] We have also obtained results that incorporate the estimated uncertainty in our parameters. Even here, 95% confidence intervals for black students, females, and students with above-median high school GPAs are all below zero, as is the interval for the effect over all students ("total"), reinforcing the finding that sorting significantly affects student achievement. The reason that females and black students lose significantly more when friendships have been homogenized is that these groups have other characteristics associated with higher estimated study types (as do their friends). For example, Table 1 shows that females have high school GPAs that are, on average, over half a standard deviation higher than males, while having similar average high school study time. On average, black students have over 40% of a standard deviation higher high school study time than nonblack students.

Figure 6: Effect of homogenizing friends on average achievement (GPA points), across simulated networks



Note: The lower and upper ends of each bar respectively denote the 2.5th and 97.5th percentile of the mean change in achievement across all simulations, for the group indicated on the x-axis. The mean over all simulations for each group (presented in Col. (3) of Table 9) is denoted by a circle.

9 Conclusion

This paper presents an equilibrium model of student study time choices and the production of achievement. Social interactions are present because costs of study time for a student depend on the study times of that student's peers. We estimate this model and provide evidence that this mechanism is important in the production of academic achievement. Our approach was made possible by three key features of the BPS: direct measurements of study time, measurements of a social network for a cohort of Berea students, and measures of student propensities to study. We develop a specification test that can detect unobserved determinants of study time.

We use the structural model to examine counterfactuals that are informative about the role of network feedback effects and sorting in peer characteristics. Heterogeneity in student characteristics and how students are interconnected determine the distribution of responses to changes in a student's study time. Our structural approach provides a very clear and intuitive interpretation for quantities of policy interest. For example, we estimate substantial best response heterogeneity, wherein the most reactive student has a best response function slope more than twice as steep as that of the least reactive student. Our results indicate that equilibrium effects, mediated by the whole social network, are quantitatively important in determining the responses of network-wide study time and achievement to shocks in study time. In addition, our results indicate that homophily, or sorting in peers' characteristics, plays an important role in the production of achievement. For example, homogenizing friends would reduce average achievement by 0.02 GPA points and the standard deviation of achievement by 5%. The results of our specification test suggest that our study propensity measures play a crucial role in addressing endogeneity concerns.

A Data

A.1 Survey questions

Figure 7: Time diary question

Survey #5 (Please complete both sides of this sheet) CPO 1971 (3)

Question A.

Reminders: Be sure to put an arrow (-->) next to the time that it is right now. And label this arrow with the words **YESTERDAY** and **START**.

Beginning with the **What were you doing** box next to the arrow, fill in your activities starting 24 hours ago (yesterday) and ending right before you began completing this survey.

Please use the 13 words listed in BOLD on the right of this page to describe your activities.

Time Period	What were you doing?	Time Period	What were you doing?
MORNING		EVENING	
6:00 AM		6:00 PM	
6:20 AM		6:20 PM	
6:40 AM		6:40 PM	
7:00 AM		7:00 PM	
7:20 AM		7:20 PM	
7:40 AM		7:40 PM	
8:00 AM		8:00 PM	
8:20 AM		8:20 PM	
8:40 AM		8:40 PM	
9:00 AM		9:00 PM	
9:20 AM		9:20 PM	
9:40 AM		9:40 PM	
10:00 AM		10:00 PM	
10:20 AM		10:20 PM	
10:40 AM		10:40 PM	
11:00 AM		11:00 PM	
11:20 AM		11:20 PM	
11:40 AM		11:40 PM	
AFTERNOON		NIGHT	
12:00 noon		12:00 midnight	
12:20 PM		12:20 AM	
12:40 PM		12:40 AM	
1:00 PM		1:00 AM	
1:20 PM		1:20 AM	
1:40 PM		1:40 AM	
2:00 PM		2:00 AM	
2:20 PM		2:20 AM	
2:40 PM		2:40 AM	
3:00 PM		3:00 AM	
3:20 PM		3:20 AM	
3:40 PM		3:40 AM	
4:00 PM		4:00 AM	
4:20 PM		4:20 AM	
4:40 PM		4:40 AM	
5:00 PM		5:00 AM	
5:20 PM		5:20 AM	
5:40 PM		5:40 AM	

LIST OF WORDS in bold

In Class
Attending class, attending labs, attending required class sessions

Studying (Outside of class time)
(refer to **pg 2** for more details)

Athletics
(Intercollegiate or Intramural - games or practice)

Clubs

Exercising

Recreation
(reading which is unrelated to courses, listening to music, watching movie, spending time with friends, etc.)

Shopping

Eating

Sleeping

Partying

Personal

Working (in Labor position)

Other
(Please describe on your sheet)

Figure 8: Friends question

(6)

Question K. Please write down the **first and last names** of the four people that have been your best friends at Berea College during **this fall term (2001)**. That is, write down the names of the four people with whom you have been spending the most time during the fall term. Also please mention how many hours per week you spend with each person and how many hours you spend studying or talking about classes with each person.
Please include your boyfriend/girlfriend or husband/wife if they are among your four best friends.
Also include your roommate if he/she is among your four best friends.
Place a check next to the name of your boyfriend/girlfriend or husband/wife.

Four best friends	**Hours spent with this person in a typical week**	**Hours spent with this person studying /talking about classes in a typical week**
1. ____________	______	______
2. ____________	______	______
3. ____________	______	______
4. ____________	______	______

Current Roommate: (Your roommate should be listed above also if he/she is one of four best friends.)

____________ ______ ______

Question L. How would you describe your relationship with your roommate during **Fall term**? **Circle one.**

1. Good friends, spent a lot of time together.
2. Got along OK, but didn't spend much time together.
3. Didn't get along very well.
4. Had significant conflicts.
5. Did not have a roommate.

Question M. 1) Since the start of the fall term, did your father lose his job without being able to find a similarly paying replacement job? **Note**: Please answer NO if your father did not lose his job, lost his job but found a similarly paying new job, did not work at a job for pay, you do not know the status of your father, or your father is deceased.

Yes **No** **Not applicable**

2) Since the start of the fall term did your mother lose her job without being able to find a similarly paying replacement job? Note: Please refer to the **Note** in part 1) of this question.

Yes **No** **Not applicable**

Question N. Have you encountered any academic difficulties during the **fall term**? Circle one. **Yes** **No**

If you have encountered academic difficulties during the **fall term**, please circle the people that you discussed these problems with? Also indicate whether each circled item was helpful or not in providing encouragement.

1. Parents	helpful	not helpful
2. Family members other than parents	helpful	not helpful
3. Friends at Berea	helpful	not helpful
4. Friends not at Berea	helpful	not helpful
5. Counselors, Advisers, or Teachers at Berea	helpful	not helpful
6. Former high school or elementary teachers	helpful	not helpful
7. Other (describe) ____________	helpful	not helpful

B Additional Model Material

B.1 Concavity of Best Response Function

The optimal choice of study time for the period game solves the function $G(s, s_{-i}) = \frac{\partial c}{\partial s} - \beta_2 = 0$. To find how s varies with friend study time, use the Implicit Function Theorem:

$$\frac{\partial s}{\partial s_{-i}} = -\frac{\frac{\partial G}{\partial s_{-i}}}{\frac{\partial G}{\partial s}} = -\frac{\frac{\partial^2 c}{\partial s \partial s_{-i}}}{\frac{\partial^2 c}{\partial s^2}}.$$

If friend study time decreases the cost of increasing one's own study time, the numerator is positive. If the cost of studying is convex in own study time, the denominator is negative, meaning the overall sign is positive. Moreover, if friend study time enters $c(\cdot)$ in a weakly concave manner, e.g., $\tau_s \leq 1$, the numerator is weakly smaller in absolute value for larger values of s_{-i}, i.e., study time is weakly concave in friend study time.

B.2 Proof of Existence and Uniqueness of Equilibrium

Claim 2. *Let k be a number strictly greater than 24. There exists a unique pure strategy Nash equilibrium if $\psi_i : R^N \mapsto R$ are weakly concave and weakly increasing, $\psi_i(0) > 0$, and $\psi_i(k) < k$ for $i \in N$.*

Proof. Define $\mathbf{S} = [0, k]^N$, i.e., a compact and convex set. Define a function Ψ:

$$\Psi : \mathbf{S} \mapsto \mathbf{S} = \begin{bmatrix} \psi_1(x_{-1}) \\ \psi_2(x_{-2}) \\ \vdots \\ \psi_N(x_{-N}) \end{bmatrix}.$$

Existence: $\Psi(\cdot)$ is a continuous self map on the compact set $\mathbf{S}$, so an equilibrium exists by Brouwer's Fixed Point Theorem.

Uniqueness: If $\Psi(\cdot)$ is strictly concave and weakly increasing we can apply Kennan (2001). Next, consider the case where $\Psi(\cdot)$ is linear, in which case we can prove $\Psi(\cdot)$ is a contraction. Write the linear form of $\Psi(\cdot)$ as

$$\Psi(X) = \begin{bmatrix} \alpha_{11} + \alpha_{21} x_{-1} \\ \alpha_{12} + \alpha_{22} x_{-2} \\ \vdots \\ \alpha_{1N} + \alpha_{2N} x_{-N} \end{bmatrix},$$

where, by assumption, $\max_{i \in N}\{\alpha_{2i}\} < 1$. Let distance be calculated according to the taxicab distance, i.e., $d(X_1, X_2) = \sum_{g \in N} |X_{1g} - X_{2g}|$ for $X_1, X_2 \in \mathbf{S}$. The Contraction Mapping Theorem holds if $d(\Psi(X_1), \Psi(X_2)) \leq bd(X_1, X_2)$, for $b \in (0, 1)$. Calculating this for the special case where Ψ is a linear map, we have

$$d(\Psi(X_1), \Psi(X_2)) = \sum_{i \in N} \alpha_{2i}|X_1 - X_2| \leq \max_{i \in N}\{\alpha_{2i}\}|X_1 - X_2| < d(X_1, X_2),$$

i.e., the condition for the Contraction Mapping Theorem is satisfied, where $b = \max_{i \in N}\{\alpha_{2i}\} \in (0, 1)$. □

B.3 Other Mechanisms for Social Interactions

B.3.1 Conformity Specification of Cost Function

We refer the cost function specification in (3) as the "cost-reduction model". Consider the alternative effort cost function, which we refer to as the "conformity model":

$$c(s, s_{-i}, \mu_{si}) = (\delta_1 + \delta_2\gamma(\mu_{si}))\, s + \frac{\delta_3}{2}s^2 + \frac{\delta_4}{2}\left(s - (1 + \delta_5\gamma(\mu_{si}))\, s_{-i}^{\tau_s}\right)^2. \tag{38}$$

Solving the student's problem results in the best response function

$$s_i = \frac{\beta_2 - \delta_1}{\delta_3 + \delta_4} - \frac{\delta_2}{\delta_3 + \delta_4}\gamma(\mu_{si}) + \frac{\delta_4}{\delta_3 + \delta_4}s_{-i}^{\tau_s} + \frac{\delta_4\delta_5}{\delta_3 + \delta_4}\gamma(\mu_{si})s_{-i}^{\tau_s}. \tag{39}$$

If we make a similar normalization as was performed for the cost-reduction model, by setting $\delta_3 = 1$, we obtain

$$s_i = \underbrace{\frac{\beta_2 - \delta_1}{1 + \delta_4}}_{-\theta_3} + \underbrace{\frac{-\delta_2}{1 + \delta_4}}_{-\theta_4}\gamma(\mu_{si}) + \underbrace{\frac{\delta_4}{1 + \delta_4}}_{(\beta_2 - \theta_1)}s_{-i}^{\tau_s} + \underbrace{\frac{\delta_4\delta_5}{1 + \delta_4}}_{-\theta_2}\gamma(\mu_{si})s_{-i}^{\tau_s}. \tag{40}$$

That is, we can represent the parameters in equation (38) above in terms of parameters in (3), which are in braces beneath their counterparts in the conformity model in (40). Therefore, the distinction between the different formulations of the cost function—a cost of deviating from friend behavior vs. a cost (reduction) from studying with ones friends—has no empirical content.

What matters, then, is interpretation and intuition of the two cost specifications. The cost-reduction model captures the intuitive notion that friends studying more, i.e., a higher s_{-i} entering the denominator in the last three terms of equation (3), would reduce the cost of studying by making it more enjoyable. At the same time, it also makes intuitive sense that there would also be a private component to the cost of studying (the first two terms in the expression). Note that τ_s also has an intuitive meaning in the cost-reduction model, as it represents the degree to which the utility gains from friends' studying diminish. If friends studying more reduces your marginal cost of studying because you enjoy spending time with them, then it would be natural to allow this benefit to decrease the more they study, i.e., $\tau_s < 1$.

Now consider the conformity model, first setting $\tau_s = 1$ and $\delta_5 = 0$. The interpretations of δ_1 and δ_2 are quite natural: students have private costs of studying, which may be heterogeneous. The interpretation of δ_4 is also quite natural: if positive, this creates a force inducing friends to behave similarly by conforming in their choice of study time.

Now consider $\tau_s < 1$ and assume $\delta_4 > 0$ (but maintain $\delta_5 = 0$). We believe the interpretation in this case for the conformity model is less clear than it is in the cost-reduction model. If $s > s_{-i}$ for a particular student then reducing τ_s would create a bigger cost of deviating. On the other hand, if $s < s_{-i}$ then reducing τ_s would decrease the cost of deviating. The interpretation of $\delta_5 \neq 0$ (maintain $\tau_s = 1$) is also slightly less clear, for similar reasons: The thing one is conforming to is not the same as one's choice: A student might choose to study four hours a day to "conform" to her friends who study five hours, not due to the private cost of studying. (Note that instead having the heterogeneous term outside the quadratic term (i.e., added to δ_4) would break the observational equivalence of the two cost functions.) Therefore, though in the homogeneous, linear model it does appear that conformity has a nice intuition, we believe the cost-reduction specification is a bit more intuitive when considering nonlinear and/or heterogeneous best response functions.

B.3.2 Production Complementarities

Suppose we did not have achievement data. For simplicity, consider the homogeneous, linear, best response specification (i.e., $\theta_2, \theta_4 = 0$); the following result also obtains when using the more general specification of the cost function. Consider the following specification of our achievement equation:

$$y(s_{it}, \mu_{yi}, s_{-it}) = \beta_1 + \beta_2 s_{it} + \beta_{3,\text{comp}} \frac{s_{it}}{s_{-it}} + \mu_{yi}, \tag{41}$$

where s_{it} is own study effort and s_{-it} is friend study effort. If $\beta_{3,\text{comp}} < 0$, then increases in peer effort increase the marginal product of one's own effort.[43] The student's problem would

[43] Note, one could instead have defined the production complementarity according to $\beta_{3,\text{comp}} s_{it} s_{-it}$, where $\beta_{3,\text{comp}} > 0$ would correspond to production complementarities. Though such a functional form would technically produce identification in the cost-reduction model we consider, it would not in the conformity model (where the interaction above, $\beta_{3,\text{comp}} s_{it}/s_{-it}$, for similar reasons, would produce identification via functional form).

still be separable across periods, resulting in the best response function

$$s_{it} = (\beta_{3,\text{comp}} - \theta_3) + (\beta_2 - \theta_1)s_{-it}. \tag{42}$$

It is obvious from (42) that we cannot separately identify $\beta_{3,\text{comp}}$ and θ_3 without having data on the marginal product of inputs (i.e., data on achievement outcomes). Indeed, this is the same argument that, without data on achievement, we could not identify the extent to which students study because it is enjoyable (θ_1) versus doing so because it affects their achievement (β_2). On the other hand, having both study time and achievement data would clearly allow one to identify the extent to which production complementarities underlie social interactions.

C Additional Estimation Results

C.1 Human Capital Spillovers ("Contextual Effects")

Table 10 presents the estimation results of our specifications allowing for contextual effects in achievement, or human capital spillovers, described in Section 7.3. Specification (1) presents the baseline estimates (i.e. those where friend characteristics do not directly affect achievement), specification (2) presents results obtained when we re-estimated parameters allowing for contextual effects generated by human capital type, as in (35), and specification (3) presents results obtained when we re-estimated parameters allowing for achievement contextual effects generated by the more flexible technology described in (36).

In specification (2), we obtain a point estimate on the achievement contextual effect parameter of $\widehat{\beta}_{3,\text{cont}} = 0.111$, which has a standard error of 0.138. Similarly, in specification (3), none of the estimated coefficients in $\omega_{y,\text{cont}}$, reported in the bottom six rows of the top panel, are significantly different from zero. The coefficient on friend HS GPA ($\omega_{y,\text{HS GPA}}$), which would seem to be the most likely source of direct achievement spillovers, is one-tenth the value of the (significant) coefficient on HS GPA in one's own human capital type and not significantly different from zero. The contextual characteristic with the most explanatory power for achievement is the share of male friends ($\omega_{y,\text{Male}}$), although this too is not a significant determinant of achievement. Based on a likelihood-ratio test, we would not reject the baseline model for that in specification (3) at any conventional significance level (the likelihood ratio test statistic has a p-value of 0.3673). We further note that there is a striking similarity between the point estimates and statistical significance of common parameters estimated under the baseline and under both specifications allowing for achievement contextual effects. This means the inclusion of such effects would not appreciably change our quantitative (or qualitative) results.

Our *a priori* belief was that our direct collection of study time data would diminish the potential role played by contextual effects. Because we do not find evidence supporting the direct transmission of peer characteristics in academic achievement, we have retained our baseline specification for the exposition of our results.

C.2 Production Complementarities

Table 11 presents the estimation results of our specification allowing for production complementarities, as described in Section 7.4. Specification (1) is the baseline estimates (i.e. those without production complementarities) and specification (2) allows for production complementarities, as in (37). Unlike the extensions including achievement contextual effects, extending the model to allow for production complementarities results in a different best response function:[44]

$$s_{it} = (\beta_{3,\text{comp}} - \theta_3) - \theta_4\gamma(\mu_{si}) + (\beta_2 - \theta_1)s_{-it} - \theta_2\gamma(\mu_{si})s_{-it}. \tag{43}$$

We obtain a point estimate on the production complementarity parameter of $\widehat{\beta}_{3,\text{comp}} = 0.904$, which has a standard error of 0.633. Although this estimate may have a surprising sign, wherein increases in friend study time *reduce* the marginal product of one's own study time (potentially due to friends goofing off when studying together), the statistically insignificant estimate does not provide evidence supporting production complementarities in achievement as the source generating social interactions in our application; rather a cost-based (or, equivalently, conformity-based) mechanism seems to generate the data. As suggested by inspection of (43), the parameter most affected by this extension is θ_3, the intercept of the best response function. Notably, parameters governing best response function slopes (θ_1, θ_2, and the study type parameters ω_s), which determine the level and distribution of the effects of social interactions in study time, are relatively unaffected.

C.3 Out-of-Sample Validation

Table 12 presents the baseline parameter estimates (Col. (1)) and those obtained when using only first-semester data (Col. (2)). The sets of parameters are strikingly similar between the two columns; this is confirmed by their having very similar (first-semester-only) log likelihoods, which are presented at the bottom of each column. This suggests that the out-of-sample fit of second-semester outcomes may be reasonable when based on parameters

[44] As in the baseline specification, we have set the best response function to be linear.

Table 10: Estimates for Contextual Effects Specifications

Parameter	Baseline Estimate	Baseline SE	Human Capital Type Contextual Estimate	Human Capital Type Contextual SE	Flexible Contextual Estimate	Flexible Contextual SE
	(1)		(2)		(3)	
Production function*						
β_1	-0.350	0.4185	-0.591	0.5419	-0.895	0.6453
β_2	0.254	0.0651	0.258	0.0681	0.332	0.0946
$\beta_{3,\text{cont}}$			0.111	0.1384		
$\omega_{y,\text{HS GPA}}$	0.470	0.0808	0.462	0.0817	0.449	0.0868
$\omega_{y,\text{ACT}}$	0.047	0.0112	0.046	0.0114	0.047	0.0119
$\omega_{y,\text{Black}}$	-0.213	0.1074	-0.182	0.1141	-0.125	0.1401
$\omega_{y,\text{Male}}$	-0.037	0.0849	-0.021	0.0822	-0.080	0.0991
$\omega_{y,\text{HS study}}$	-0.007	0.0042	-0.007	0.0044	-0.009	0.0053
$\omega_{y,\text{expected study}}$	-0.005	0.0035	-0.005	0.0035	-0.005	0.0037
$\omega_{y,\text{cont,HS GPA}}$					0.044	0.1251
$\omega_{y,\text{cont,ACT}}$					0.005	0.0157
$\omega_{y,\text{cont,Black}}$					-0.119	0.1747
$\omega_{y,\text{cont,Male}}$					0.189	0.1231
$\omega_{y,\text{cont,HS study}}$					-0.010	0.0073
$\omega_{y,\text{cont,expected study}}$					0.006	0.0050
Study cost function / Best response function**						
θ_1	-1.074	0.1551	-1.068	0.1569	-1.051	0.1563
θ_2	0.874	0.2351	0.870	0.2355	0.899	0.2063
θ_3	-0.907	0.8097	-0.918	0.8084	-0.981	0.8805
θ_4	0.096	1.2800	0.098	1.2713	0.183	1.2996
$\tau_{\mu,1}$	0.105	0.0601	0.105	0.0604	0.089	0.0479
$\tau_{\mu,2}$	-0.003	0.0028	-0.003	0.0028	-0.003	0.0023
$\omega_{s,\text{HS GPA}}$***	1.000	–	1.000	–	1.000	–
$\omega_{s,\text{ACT}}$	-0.063	0.0870	-0.065	0.0890	-0.063	0.0930
$\omega_{s,\text{Black}}$	-0.735	0.7459	-0.720	0.7516	-0.710	0.7782
$\omega_{s,\text{Male}}$	-1.065	0.7892	-1.081	0.8004	-1.188	0.8562
$\omega_{s,\text{HS study}}$	0.344	0.1554	0.347	0.1578	0.350	0.1651
$\omega_{s,\text{expected study}}$	0.005	0.0309	0.004	0.0311	-0.0003	0.0324
Shocks						
σ_ϵ	0.721	0.0185	0.721	0.0185	0.717	0.0187
σ_η	2.159	0.0377	2.159	0.0377	2.159	0.0376
Log Likelihood	-4696.361		-4695.932		-4693.100	

* Production function in specifications (1)-(2) is $y = \beta_1 + \beta_2 s_{it} + \beta_{3,\text{cont}}\mu_{y,-i,t} + \mu_{yi}$, where $\mu_{yi} = x_i'\omega_y$ and $\mu_{y,-i,t} \equiv \frac{\sum_{j=1}^N A_t(i,j)\mu_{yj}}{\sum_{j=1}^N A_t(i,j)}$. Production function in specification (3) is $y = \beta_1 + \beta_2 s_{it} + \mu_{yi} + \mu_{y,\text{cont},-i,t}$, where $\mu_{yi} = x_i'\omega_y$, $\mu_{y,\text{cont},-i,t} \equiv \frac{\sum_{j=1}^N A_t(i,j)\mu_{y,\text{cont},j}}{\sum_{j=1}^N A_t(i,j)}$, and $\mu_{y,\text{cont},j} = x_j'\omega_{y,\text{cont}}$.

** Best response function: $s_{it} = -\theta_3 - \theta_4\gamma(\mu_{si}) + (\beta_2 - \theta_1)s_{-it} - \theta_2\gamma(\mu_{si})s_{-it}$, where $\mu_{si} = x_i'\omega_s$. As in the baseline estimates, we have set $\tau_s = 1$.

***: Normalized to 1.

Table 11: Estimates for Production Complementarities Specification

Parameter	Baseline Estimate	SE	Prod. Complementarities Estimate	SE
	(1)		(2)	
Production function*				
β_1	-0.350	0.4185	-1.099	0.6830
β_2	0.254	0.0651	0.267	0.0693
$\beta_{3,\text{comp}}$			0.904	0.6334
$\omega_{y,\text{HS GPA}}$	0.470	0.0808	0.430	0.0956
$\omega_{y,\text{ACT}}$	0.047	0.0112	0.048	0.0129
$\omega_{y,\text{Black}}$	-0.213	0.1074	-0.191	0.1195
$\omega_{y,\text{Male}}$	-0.037	0.0849	-0.017	0.0928
$\omega_{y,\text{HS study}}$	-0.007	0.0042	-0.013	0.0061
$\omega_{y,\text{expected study}}$	-0.005	0.0035	-0.006	0.0040
Study cost function / Best response function**				
θ_1	-1.074	0.1551	-1.056	0.1495
θ_2	0.874	0.2351	0.839	0.2256
θ_3	-0.907	0.8097	-0.251	1.2293
θ_4	0.096	1.2800	0.561	1.4312
$\tau_{\mu,1}$	0.105	0.0601	0.108	0.0595
$\tau_{\mu,2}$	-0.003	0.0028	-0.004	0.0028
$\omega_{s,\text{HS GPA}}$***	1.000	–	1.000	–
$\omega_{s,\text{ACT}}$	-0.063	0.0870	-0.055	0.0760
$\omega_{s,\text{Black}}$	-0.735	0.7459	-0.585	0.6456
$\omega_{s,\text{Male}}$	-1.065	0.7892	-0.886	0.6592
$\omega_{s,\text{HS study}}$	0.344	0.1554	0.297	0.1284
$\omega_{s,\text{expected study}}$	0.005	0.0309	0.009	0.0273
Shocks				
σ_ϵ	0.721	0.0185	0.719	0.0183
σ_η	2.159	0.0377	2.160	0.0377
Log Likelihood:	-4696.361		-4694.625	

* Production function: $y = \beta_1 + \beta_2 s_{it} + \beta_{3,\text{comp}} \frac{s_{it}}{s_{-it}} + \mu_{yi}$, where $\mu_{yi} = x_i'\omega_y$.

** Best response function: $s_{it} = (\beta_{3,\text{comp}} - \theta_3) - \theta_4\gamma(\mu_{si}) + (\beta_2 - \theta_1)s_{-it} - \theta_2\gamma(\mu_{si})s_{-it}$, where $\mu_{si} = x_i'\omega_s$. As in the baseline estimates, we have set $\tau_s = 1$.

***: Normalized to 1.

estimated using only first-semester data.

Table 12: Parameters Under Baseline and Only-First-Semester Data

	Baseline	Only First Semester
	(1)	(2)
β_1	-0.350	-0.154
β_2	0.254	0.230
$\omega_{y,\text{HS GPA}}$	0.470	0.400
$\omega_{y,\text{ACT}}$	0.047	0.053
$\omega_{y,\text{Black}}$	-0.213	-0.278
$\omega_{y,\text{Male}}$	-0.037	-0.035
$\omega_{y,\text{HS study}}$	-0.007	-0.009
$\omega_{y,\text{expected study}}$	-0.005	-0.005
$\omega^*_{s,\text{HS GPA}}$	1.000	1.000
$\omega_{s,\text{ACT}}$	-0.063	-0.121
$\omega_{s,\text{Black}}$	-0.735	-0.765
$\omega_{s,\text{Male}}$	-1.065	-0.873
$\omega_{s,\text{HS study}}$	0.344	0.271
$\omega_{s,\text{expected study}}$	0.005	0.000
σ_η	2.159	2.203
σ_ϵ	0.721	0.709
τ^{**}_s	1.000	1.000
$\tau_{\mu,1}$	0.105	0.127
$\tau_{\mu,2}$	-0.003	-0.005
θ_1	-1.074	-1.103
θ_2	0.874	0.848
θ_3	-0.907	-1.138
θ_4	0.096	0.144
Log Likelihood: (first-semester)	-2375.58255	-2373.61342

Note: *: Normalized to 1. **: Fixed to 1. Col. (1) presents estimates from baseline model and the log likelihood in the first semester. Col. (2) presents results when parameters were estimated using only first-semester data and the log likelihood in the first semester.

To further investigate, we compare model fit for the first-semester data, which was used to estimate model parameters, and the validation data from the second semester. We can see that the out-of-sample fit for study time (Figure 9), GPA (Figure 10), and own-vs.-friend study time (Figure 11) all seem quite good.

D Additional Tables

Figure 9: In and-Out-of-sample fit; study time

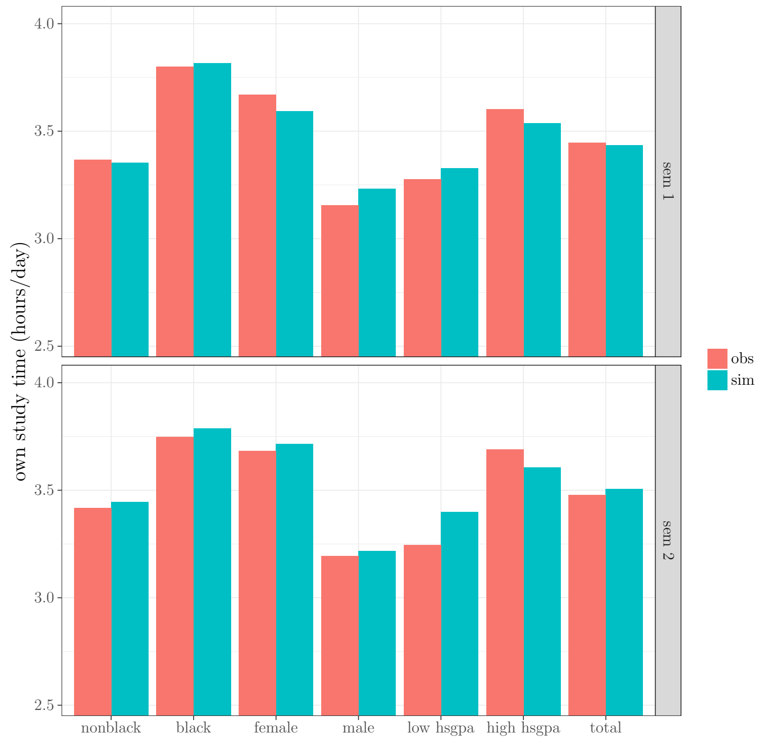


Figure 10: In- and-Out-of-sample fit; GPA

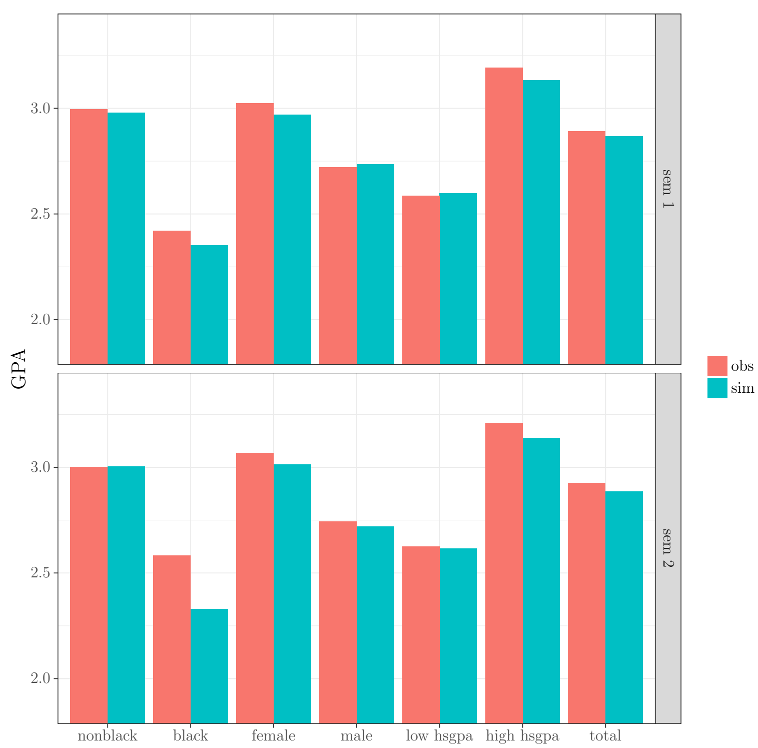


Figure 11: In- and-Out-of-sample fit; own vs. friend study time

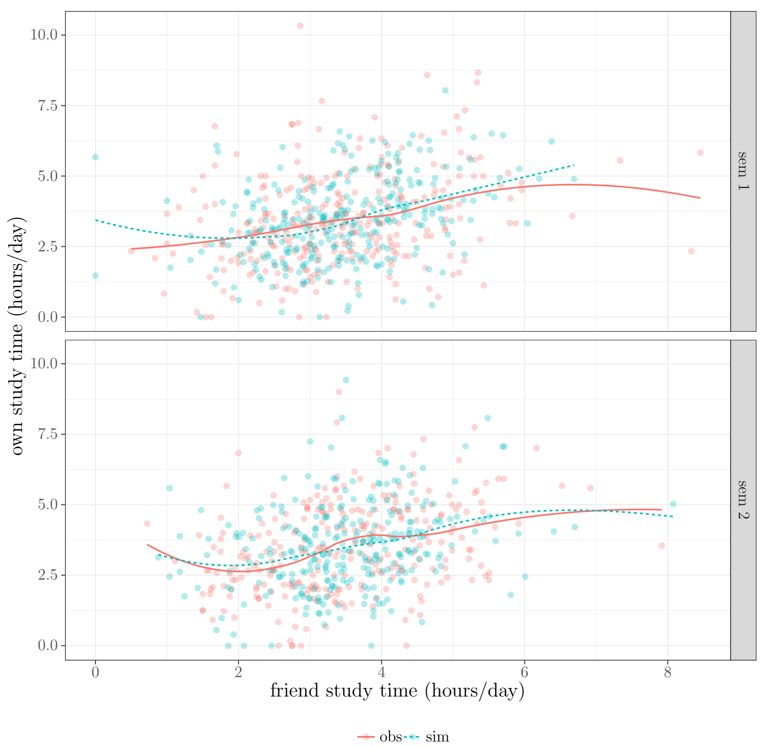


Table 13: Study time regressions controlling for different sets of characteristics, pooled over both semesters

	Dependent variable: Own study			
	(1)	(2)	(3)	(4)
Male	−0.369***	−0.328**	−0.391***	
	(0.136)	(0.140)	(0.135)	
Black	0.116	0.333*	0.324*	
	(0.186)	(0.192)	(0.172)	
HS GPA	0.413***	0.392**		
	(0.149)	(0.156)		
ACT	−0.032	−0.029		
	(0.021)	(0.022)		
HS study	0.043***			
	(0.006)			
Expected study	−0.002			
	(0.006)			
Friends study	0.166***	0.198***	0.202***	0.228***
	(0.037)	(0.039)	(0.039)	(0.038)
Constant	1.915***	2.167***	2.850***	2.648***
	(0.671)	(0.679)	(0.172)	(0.152)
Observations	574	574	574	574
R^2	0.169	0.087	0.076	0.058

Note: *p<0.1; **p<0.05; ***p<0.01 GPA is measured in GPA points (0-4). HS study and expected study are measured in hours/week. Own and friend study are measured in hours/day. The variable "Friend z" for student i in period t is the average of the variable z across i's friends in period t.

E Dyadic, Separate Networks with Homogeneous Best Responses

To make the test statistic more concrete, in this section we develop an example environment with dyadic, separate networks with homogeneous best responses.

Table 14: Study time regressions, pooled over both semesters

	Dependent variable: Own study	
	(1)	(2)
Male	−0.369***	−0.365***
	(0.136)	(0.137)
Black	0.116	0.115
	(0.186)	(0.187)
HS GPA	0.413***	0.389***
	(0.149)	(0.150)
ACT	−0.032	−0.034
	(0.021)	(0.021)
HS study	0.043***	0.041***
	(0.006)	(0.006)
Expected study	−0.002	−0.002
	(0.006)	(0.006)
Own share science courses		0.349
		(0.390)
Friend study	0.166***	0.157***
	(0.037)	(0.038)
Avg. friend share science courses		0.880
		(0.562)
Constant	1.915***	1.873***
	(0.671)	(0.684)
Observations	574	574
R^2	0.169	0.176

Note: *p<0.1; **p<0.05; ***p<0.01 GPA is measured in GPA points (0-4). Own and friend HS study and expected study are measured in hours/week. Own and friend study are measured in hours/day. The variable "Friend z" for student i in period t is the average of the variable z across i's friends in period t.

Example 1 (Dyadic, separate networks with homogeneous best responses)**.** *Consider the following special case, with*

$$F_1(\Gamma_2) = X\lambda_1 \tag{44}$$

$$W_t(X;\Gamma_2) = W\lambda_2 \tag{45}$$

$$\Pi(X) = X\lambda_3, \tag{46}$$

where λ_1, λ_2 *are scalars,* λ_3 *is a matrix of parameters, and* $W = I_N \otimes I_2$ *is a block diagonal matrix with* I_2 *(i.e., two-by-two Identity matrices) along the diagonal, representing the fact that each student is friends with exactly one other student. Then (31) would become*

$$\widetilde{\epsilon} \approx \{(I - W\widetilde{\lambda}_2)^{-1}[X(\lambda_1 + \lambda_3)] - (I - W\widetilde{\lambda}_2)^{-1}[X\widehat{\lambda}_1]\} + (I - W\lambda_2)^{-1}u + \eta_s. \tag{47}$$

Note that plim $\widehat{\lambda}_1 = (\lambda_1 + \lambda_3)$*, i.e., where will be no prediction bias, meaning we are in Case (i) above. Because* W *is block diagonal, it is sufficient to consider the top-left* 2×2 *block, representing the first friendship dyad. The expression (47) for these students then, eliminating the* η_s*, which are independently distributed from all other variables and therefore qualitatively immaterial in the following calculation, is*

$$\begin{bmatrix} \widetilde{\epsilon}_1 \\ \widetilde{\epsilon}_2 \end{bmatrix} \approx \frac{1}{1-\lambda_2^2} \begin{bmatrix} 1 & \lambda_2 \\ \lambda_2 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \frac{1}{1-\lambda_2^2} \begin{bmatrix} u_1 + \lambda_2 u_2 \\ \lambda_2 u_1 + u_2 \end{bmatrix}, \tag{48}$$

giving the product of residuals for students in the first dyad: $\left(\frac{1}{1-\lambda_2^2}\right)^2 (\lambda_2 u_1^2 + (1+\lambda_2^2)u_1u_2 + \lambda_2 u_2^2)$. *Under the maintained assumption that* $\lambda \geq 0$*, the expectation of this product can only be zero if* $u_1 u_2 < 0$*, i.e., own and friend errors* u *are negatively correlated, and in exactly the right way.*

References

Badev, A., "Discrete Games in Endogenous Networks: Theory and Policy," *Working Paper*, 2013.

Ballester, C., A. Calvó-Armengol and Y. Zenou, "Who's Who in Networks. Wanted: The Key Player," *Econometrica*, 74(5):1403–1417, 2006.

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

Boucher, V. and B. Fortin, "Some Challenges in the Empirics of the Effects of Networks," in

Y. Bramoullé, A. Galeotti and B. Rogers, eds., "Handbook on the Economics of Networks," chap. 12, Oxford University Press, Oxford, 2016.

Bramoullé, Y., S. Currarini, M. Jackson, P. Pin and B. W. Rogers, "Homophily and Long-Run Integration in Social Networks," *Journal of Economic Theory*, 147(5):1754–1786, 2012.

Brock, W. A. and S. N. Durlauf, "Interactions-Based Models." in J. J. Heckman and E. E. Leamer, eds., "Handbook of Econometrics," vol. 5, pp. 235–260, Amsterdam: North-Holland Press, 2001.

Bursztyn, L., F. Ederer, B. Ferman and N. Yuchtman, "Understanding Mechanisms Underlying Peer Effects: Evidence From a Field Experiment On Financial Decisions," *Econometrica*, 82(4):1273–1301, 2014.

Calvó-Armengol, A., E. Patacchini and Y. Zenou, "Peer Effects and Social Networks in Education," *The Review of Economic Studies*, 76(4):1239–1267, 2009.

Camargo, B., R. Stinebrickner and T. R. Stinebrickner, "Interracial Friendships in College," *Journal of Labor Economics*, 28(4):861–892, 2010.

Christakis, N., J. Fowler, G. Imbens and K. Kalyanaraman, "An Empirical Model for Strategic Network Formation," *NBER Working Paper No. 16039*, 2010.

Cooley Fruehwirth, J., "Identifying Peer Achievement Spillovers: Implications for Desegregation and the Achievement Gap," *Quantitative Economics*, 4(1):85–124, 2013.

Csardi, G. and T. Nepusz, "The Igraph Software Package for Complex Network Research," *InterJournal*, Complex Systems:1695, 2006.

Currarini, S., M. O. Jackson and P. Pin, "An Economic Model of Friendship: Homophily, Minorities and Segregation," *Econometrica*, 77:1003–1045, 2009.

—, "Identifying the Roles of Race-Based Choice and Chance in High School Friendship Network Formation," *Proceedings of the National Academy of Sciences*, 107(11):4857–4861, 2010.

De Giorgi, G. and M. Pellizzari, "Understanding Social Interactions: Evidence From the Classroom," *The Economic Journal*, 124(579):917–953, 2014.

de Paula, Á., S. O. Richards-Shubik and E. Tamer, "Identification of Preferences in Network Formation Games," *Working Paper*, 2016.

Epple, D. and R. Romano, "Peer Effects in Education: A Survey of the Theory and Evidence," in A. B. Jess Benhabib and M. O. Jackson, eds., "Handbook of Social Economics," vol. 1, pp. 1053–1163, North-Holland, 2011.

Freeman, L. C., "Centrality in Social Networks: Conceptual Clarification," *Social Networks*, 1(3):215–239, 1979.

Fryer Jr, R. G., "Financial Incentives and Student Achievement: Evidence From Randomized Trials," *The Quarterly Journal of Economics*, 126(4):1755–1798, 2011.

Goldsmith-Pinkham, P. and G. W. Imbens, "Social Networks and the Identification of Peer Effects," *Journal of Business & Economic Statistics*, 31(3):253–264, 2013.

Golub, B. and M. O. Jackson, "How Homophily Affects the Speed of Learning and Best-Response Dynamics," *The Quarterly Journal of Economics*, 127(3):1287–1338, 2012.

Hanushek, E. A., "Conceptual and Empirical Issues in the Estimation of Educational Production Functions," *Journal of Human Resources*, pp. 351–388, 1979.

Hsieh, C.-S. and L. F. Lee, "A Social Interactions Model with Endogenous Friendship Formation and Selectivity," *Journal of Applied Econometrics*, 31(2):301–319, 2016.

Jackson, M. O., *Social and Economic Networks*, Princeton University Press, 2008.

Jackson, M. O. and L. Yariv, "Diffusion, Strategic Interaction, and Social Structure," *Handbook of Social Economics*, 1:645–678, 2011.

Kennan, J., "Uniqueness of Positive Fixed Points for Increasing Concave Functions On N: An Elementary Result," *Review of Economic Dynamics*, 4(4):893–899, 2001.

Kline, B. and E. Tamer, "Some Interpretation of the Linear-In-Means Model of Social Interactions," *Working Paper*, 2011.

Lee, L.-F., "Asymptotic Distributions of Quasi-Maximum Likelihood Estimators for Spatial Autoregressive Models," *Econometrica*, 72(6):1899–1925, 2004.

Manski, C. F., "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 60(3):531–542, 1993.

—, "Economic Analysis of Social Interactions," *Journal of Economic Perspectives*, 14(3):115–136, 2000.

Mele, A., "A Structural Model of Segregation in Social Networks," *Working paper*, 2013.

Moffitt, R., "Policy Interventions, Low-Level Equilibria, and Social Interactions," *Social Dynamics*, 4:45–82, 2001.

Pinkse, J., M. E. Slade and C. Brett, "Spatial Price Competition: A Semiparametric Approach," *Econometrica*, 70(3):1111–1153, 2002.

R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2015.

Richards-Shubik, S. O., "Peer Effects in Sexual Initiation: Separating Demand and Supply Mechanisms," *Quantitative Economics*, 6(3):663–702, 2015, ISSN 1759-7331, doi: 10.3982/QE249.

Sacerdote, B., "Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?" in E. Hanushek, S. Machin and L. Woessmann, eds., "Handbook of the Economics of Education," vol. 3, chap. 4, pp. 249–277, Elsevier, 1st edn., 2011.

Sheng, S., "A Structural Econometric Analysis of Network Formation Games," *Working Paper*, 2014.

Stinebrickner, R. and T. R. Stinebrickner, "Time-Use and College Outcomes," *Journal of Econometrics*, 121(1):243–269, 2004.

—, "What Can Be Learned About Peer Effects Using College Roommates? Evidence From New Survey Data and Students From Disadvantaged Backgrounds," *Journal of Public Economics*, 90(8):1435–1454, 2006.

—, "The Causal Effect of Studying On Academic Performance," *The BE Journal of Economic Analysis & Policy*, 8(1):1–55, 2008a.

—, "The Effect of Credit Constraints On the College Drop-Out Decision: A Direct Approach Using a New Panel Study," *The American Economic Review*, 98(5):2163–84, 2008b.

—, "A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout," *Review of Economic Studies*, pp. 791–824, 2013.

Tincani, M. M., "Heterogeneous Peer Effects and Rank Concerns: Theory and Evidence," *Working Paper*, 2016.

Todd, P. E. and K. I. Wolpin, "On the Specification and Estimation of the Production Function for Cognitive Achievement," *The Economic Journal*, 113(485):F3–F33, 2003.

Wickham, H., *Ggplot2: Elegant Graphics for Data Analysis*, Springer New York, 2009, ISBN 978-0-387-98140-6.

Zeitlin, A., "Identification and Estimation of Peer Effects On Endogenous Affiliation Networks: An Application to Ghanaian Agriculture," Tech. rep., Working paper, 2011.