How social interactions determine input choices and outcomes in equilibrium:

Evidence from a model of study time and academic achievement

Tim Conley, Nirav Mehta, Ralph Stinebrickner, Todd Stinebrickner

February 15, 2015

PRELIMINARY (WE REALLY MEAN IT!) PLEASE DO NOT CITE

Abstract

Due to data limitations, few papers documenting the existence of peer effects explore the mechanisms through which they operate. This paper develops and estimates an equilibrium model of study time choices made by students, given a social network. The model is designed to exploit unique data collected in the Berea Panel Study (BPS). The study time data allow us to quantify an intuitive mechanism for social interactions: one's own study time may depend on friends' study time. The detailed social network data allow us to embed these individual study time choices in an equilibrium framework, allowing for feedback effects. We find evidence of both the direct and indirect effects of friend study time. Not taking into account equilibrium behavior would understate the effect of peers on achievement by 20-30% of a standard deviation.

1 Introduction

Peer effects between students are widely believed to be important for determining academic achievement. Most of the existing research on peer effects in this context has focused on establishing a causal link between academic outcomes and peer characteristics in an effort to provide evidence on whether peers matter. [1] Empirical evidence on specific mechanisms generating peer effects is very limited. In this paper we exploit unique, new data on college students from the Berea Panel Study (BPS) to provide some of the first evidence about mechanisms underlying peer effects in determining academic outcomes. We focus on a specific, intuitively appealing mechanism that a students peers study effort influences the costs of studying. We estimate an equilibrium model of study time choice conditional on a network of social contacts.

Empirical study of peer effects is difficult. A fundamental issue is that the researcher requires an operational definition of the relevant peers influencing a student. Once this has been accomplished the researcher is confronted with the question of how to identify peer influences and the mechanisms through which they operate.[2] In our context, individuals may make similar decisions about input choices due to having similar preferences for studying and/or shocks to its costs that appear to be peer effects due to sorting among friends defining our peer network.

We address the question of peer effect identification by focusing on one of the most salient potential endogeneity problems: unobserved shocks or preferences for studying may induce correlations between students' outcomes that look like peer effects. Our solution relies on data in the BPS that were intentionally collected to mitigate this issue. For tractability, we do not address the potential endogeneity of the peer network itself in this paper. We condition on an observed network and model students study time decisions in equilibrium given the network. We view this paper as a building block of a more complete model that will incorporate a model of friendship choice.

[1] See Epple and Romano (2011) for a survey of this extensive literature.

[2] see e.g. (Manski 1993), Moffit (2001), Bramouille, Djebbari, and Fortin (2009)

Peer definitions are commonly taken to be other students within an administrative unit like classroom, dormitory, or platoon. Often this is due to a lack of data on links between individuals rather than an argument that peer effects should respect these administrative unit boundaries. This peer definition can be useful when the policy question of interest regards an administrative choice regarding composition of these units, e.g. classroom tracking or platoon composition (see Gamoran [1992] or Carrell, Fullerton and West (2009)). Even if peer influences don't really respect these unit boundaries, many counterfactuals of interest can still be addressed with administrative unit peer definition. However, uncovering mechanisms underlying peer effects is difficult with such peer definitions. In our context of college students' study decisions the relevant network is sure to extend beyond classroom or dorm room and we need survey-generated data on peer connections. The BPS provides extensive data on friendship patterns among college students which enables a good operational definition of students' peer network. The BPS data contain four measurements of each student's closest friends at Berea College per year.

In addition to this direct data on network connections, the data contain detailed time-diary information. This allows us information on how much time individuals spend studying, Study effort is one of main inputs to producing human capital and arguably the most important variable input for the college students in our data. Most studies of peer effects do not have access to such data on input choices, inherently limiting their ability to uncover mechanisms underlying these effects.

Our focus on study decisions in equilibrium allows us to understand the relative importance of peer-related feedback effects. Not only might student i's outcome be influenced by i's friends, but i's friends' outcomes may be influenced by i. Moreover, these types of feedback or equilibrium effects can work indirectly through students in the social network that are more than one link away from student i. Studying such phenomena requires that one have a good measurement of the social network and possesses an empirical framework that takes feedback effects into account.

Our main preliminary results are a characterization of an impulse response to an individual increasing his or her study time and an assessment of the importance of sorting in the social network upon outcomes. An increase in an individual's study effort impacts his or her own achievement and that of all others connected in the network. Our results indicate that the effect of a one standard deviation positive impulse in study effort for a focal student upon the student's own achievement is an increase of approximately a 22% of a standard deviation. The corresponding achievement responses for students one link away from the focal student is approximately 18% of a standard deviation. The effect upon those two links from the focal student is approximately 14% of a standard deviation increase. The contribution of equilibrium feedback in these responses is substantial, it amounts to 11% of the total effect for those one link away and 28% for those two links away.

The remainder of this paper is organized as follows. Section 2 contains a literature review and description of the BPS data. Section 3 presents our model. Section 4 presents our empirical specification and operational definition of networks. Section 5 discusses our results and counterfactual exercises and once we have a conclusion it will be in Section 6.

2 Data

The Berea Panel Study (BPS) is a longitudinal survey that was designed to provide detailed information about students' input choices and educational outcomes in college, and labor market outcomes shortly after college.[3] The BPS project closely followed all students who entered Berea College in the fall of 2000 and the fall of 2001 from the time they entered college through the early portion of their working lives. Berea College is unique in certain respects that have been discussed in previous work. For example, the school has a focus on providing educational opportunities to students from relatively low income backgrounds. Therefore, caution is appropriate when considering whether results from our case study would generalize to other demographic groups or to other specific institutions. However, Berea has much in common with many colleges. It operates under a standard liberal arts curriculum and the students at Berea are similar in academic quality to, for example, students at the

[3]The BPS was designed and administered by Todd Stinebrickner and Ralph Stinebrickner.

University of Kentucky (Stinebrickner and Stinebrickner [2008]). In addition, earlier work found that academic decisions at Berea look very similar to decisions made elsewhere. For example, dropout rates are similar to those found elsewhere (for students from similar income backgrounds) and patterns of major-switching at Berea are similar in spirit to those found in a representative dataset by Arcidiacono [2004].

It is also worth noting that the study of one school is beneficial for at least two reasons. This makes it feasible to collect the multiple surveys each year that provide data that are more detailed in important dimensions than other existing surveys, Furthermore, it makes outcomes such as college grade performance directly comparable across individuals.

This paper is made possible by three types of information that are available in the BPS. First, four times each year, the BPS elicited each student's closest friends. Our main 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 similar. Second, eight times each year, the BPS collected detailed time-use information using a twenty-four hour time diary which is shown in Appendix A.1. Finally, questions on the baseline survey reveal the amount that a student studied per week in high school and the amount the person expects to study per week in college. The survey data is merged with detailed administrative data that provides information about race, gender, high school grade point average college entrance exam scores, the courses a student took in each semester, and college grade point average in each semester.

This paper focuses on the freshman year for students in the 2001 entering cohort. We focus on understanding grade outcomes during students' freshmen years for several reasons. Students have very similar course loads in their first year. Approximately 88% of all entering students in the 2001 cohort completed our baseline survey (which was administered immediately before the start of freshman classes), and response rates remained high for the eleven subsequent surveys that were administered during the freshman year. Finally, we have the most complete network data for these students and their friendship network is overwhelmingly within-cohort. Over 80% of friends reported by students in this cohort are themselves in the 2001 cohort (the large majority of the remainder are in the 2000 cohort for which we also have data). These advantages tend to fade in subsequent years as friendships change (in part due to substantial dropout after first year) and students' programs of study specialize.

2.1 Descriptive statistics

Table 1 shows summaries of student characteristics pooled over both semesters, broken down by sex, whether or not the student is black, and high school GPA level (whether or not the student's high school GPA was above the sample median). Each panel features a variable (e.g. "Male") in the top row, and then summaries of that variable for subgroups in rows beneath. For example, the second row of the first panel shows that there are 55 black students in the sample, 45% of whom are male. There are 614 observations in the sample, two semesters each for 307 students. Less than half of the sample is male and 18% are Black. The first two panels show that students with high (i.e. above-sample-median) high school GPAs are less likely to be male and black. The third panel is redundant, showing that male and black students have lower high school GPAs on average. Male and Black students have lower combined ACT scores, opposite to students with above-median high school GPAs. Black students studied more than average in high school. The bottom two panels show outcomes during the first year of college. Male students have lower average study times, while black students and students with above-median high school GPAs study more. Finally, the seventh panel shows that males, black students, and students with below-median high school GPAs all have lower average outcome GPAs than their counterparts.

Figure 1 shows the group means, with 95% confidence intervals.

Figure 1: Avg. own study time by group,with 95% confidence intervals

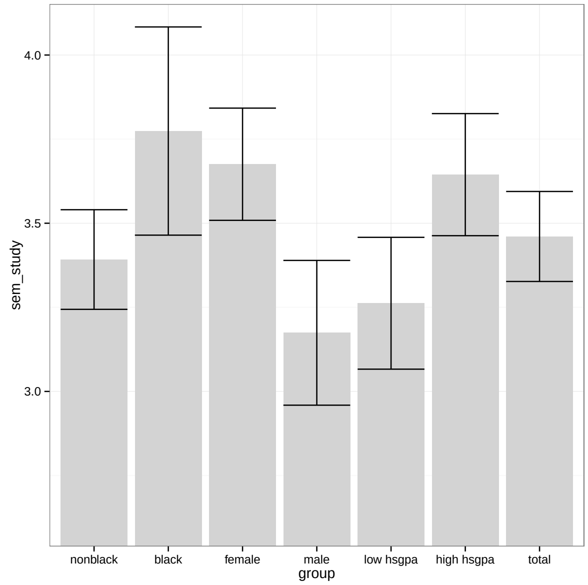


Table 2 summarizes friend data, for those who report having at least one friend in each semester, broken down by the same characteristics as in Table 1. The top panel shows that students have about 3 friends on average. The mean masks considerable variation: the minimum number of friends is one, while the maximum is 10 friends. The second and third panels show that male and black students sort towards students with the same characteristics. In addition, the third panel shows that students with above-median high school GPAs have fewer black friends. The fourth and fifth panels shows that male and black students have friends with lower incoming GPAs and combined ACT scores. The sixth and seventh panels show that black students and students with above-median high school GPAs have friends who were more studious in both high school and college, while males have less studious friends.

Table 1: Own summary statistics

Variable	Group	N	Mean	SD	Min	q1	q2	q3	Max
Male indicator	all	307	0.44	0.5	0	0	0	1	1
	given Black	55	0.45	0.5	0	0	0	1	1
	given above med. HS GPA	155	0.33	0.47	0	0	0	1	1
Black indicator	all	307	0.18	0.38	0	0	0	0	1
	given Male	134	0.19	0.39	0	0	0	0	1
	given above med. HS GPA	155	0.1	0.31	0	0	0	0	1
HS GPA	all	307	3.39	0.47	1.68	3.09	3.5	3.8	4
	given Male	134	3.24	0.51	1.68	2.9	3.21	3.7	4
	given Black	55	3.14	0.46	2.24	2.78	3.1	3.52	4
	given above med. HS GPA	155	3.77	0.17	3.5	3.6	3.8	3.9	4
ACT	all	307	23.26	3.61	14	21	23	26	33
	given Male	134	22.54	3.77	14	20	23	25	31
	given Black	55	19.91	2.51	14	18	20	21	25
	given above med. HS GPA	155	24.45	3.53	17	22	25	27	33
HS study*	all	307	11.24	11.35	0	4	8	15	70
	given Male	134	11.43	11.94	0	3.12	8	15	70
	given Black	55	15.29	14	0	5	10.5	20	70
	given above med. HS GPA	155	10.66	10.44	0	4	8	14.5	70
Sem. 1 Own study**	all	296	3.45	1.67	0	2.33	3.29	4.5	10.33
	given Male	129	3.16	1.76	0	2	2.92	4.09	8.66
	given Black	53	3.8	1.63	0	2.84	3.67	4.78	8.33
	given above med. HS GPA	153	3.6	1.7	0	2.41	3.34	4.75	8.58
Sem. 2 Own study**	all	278	3.48	1.6	0	2.23	3.34	4.75	9
	given Male	117	3.19	1.67	0	2	3	4.34	9
	given Black	50	3.75	1.55	0	2.74	3.5	4.96	7.33
	given above med. HS GPA	145	3.69	1.47	0	2.66	3.67	4.83	7.92
Sem. 1 GPA	all	307	2.89	0.78	0	2.49	3.06	3.46	4
	given Male	134	2.72	0.8	0.3	2.17	2.8	3.29	4
	given Black	55	2.42	0.78	0	1.82	2.57	2.84	4
	given above med. HS GPA	155	3.19	0.62	0.52	2.81	3.29	3.69	4
Sem. 2 GPA	all	301	2.93	0.78	0	2.53	3.05	3.46	4
	given Male	131	2.74	0.84	0	2.38	2.82	3.33	4
	given Black	53	2.58	0.86	0.44	2.22	2.62	3.33	3.78
	given above med. HS GPA	155	3.21	0.66	0	2.82	3.36	3.74	4

*Note: Hours per day spent studying senior year of high school.
**Note: Own study time is semester-level average hours per day.

Table 2: Friend summary statistics, pooled over both semesters

Variable	Group	N	Mean	SD	Min	q1	q2	q3	Max
Num. friends	all	614	3.31	1.58	1	2	3	4	10
	given Male	268	3.22	1.59	1	2	3	4	10
	given Black	110	3.21	1.35	1	2	3	4	7
	given above med. HS GPA	310	3.34	1.62	1	2	3	4	10
Friend Male	all	614	0.43	0.39	0	0	0.33	0.75	1
	given Male	268	0.74	0.31	0	0.5	0.82	1	1
	given Black	110	0.43	0.4	0	0	0.33	0.83	1
	given above med. HS GPA	310	0.35	0.38	0	0	0.25	0.67	1
Friend Black	all	614	0.18	0.32	0	0	0	0.25	1
	given Male	268	0.18	0.32	0	0	0	0.25	1
	given Black	110	0.69	0.38	0	0.43	1	1	1
	given above med. HS GPA	310	0.1	0.22	0	0	0	0	1
Friend HS GPA	all	614	3.37	0.32	2.24	3.2	3.41	3.62	4
	given Male	268	3.29	0.33	2.25	3.07	3.34	3.53	4
	given Black	110	3.18	0.34	2.25	2.96	3.19	3.41	4
	given above med. HS GPA	310	3.46	0.27	2.65	3.29	3.46	3.63	4
Friend ACT	all	614	23.29	2.63	16	21.67	23.33	25	32
	given Male	268	22.72	2.64	16.33	21	23	24.64	31
	given Black	110	21.2	2.53	16	19.33	21	22.5	29
	given above med. HS GPA	310	23.79	2.42	17.5	22.23	23.67	25.33	32
Friend HS study	all	614	11.03	7.64	0	6	9.5	14.47	70
	given Male	268	10.53	7.37	0.5	5.17	9	14	37.33
	given Black	110	14.62	7.31	2.5	9.18	13.92	18.75	37
	given above med. HS GPA	310	11.48	8.44	0.5	6	9.7	14	70
Friend study*	all	614	3.5	1.72	0	2.47	3.26	4.28	11.93
	given Male	268	3.16	1.49	0.5	2.21	3	3.88	8.46
	given Black	110	3.78	1.77	0.5	2.7	3.52	4.47	10.81
	given above med. HS GPA	310	3.64	1.79	0	2.56	3.36	4.41	11.93

*Note: Friend study time is the average of friends' own study time that semester.

Figure 2 shows the group means, with 95% confidence intervals for each group mean.

Table 3 shows other network characteristics. Both the probability a first semester friendship is not reported in the second or a second-semester friendship was not reported in the first semester are both about 50%. As demonstrated in Table 2, there is considerable sorting on race, sex, high school GPA, combined ACT score, and high school study time.

Table 4 presents OLS regression results predicting own study time (left column) and

Figure 2: Avg. friend study time by group, with 95% confidence intervals

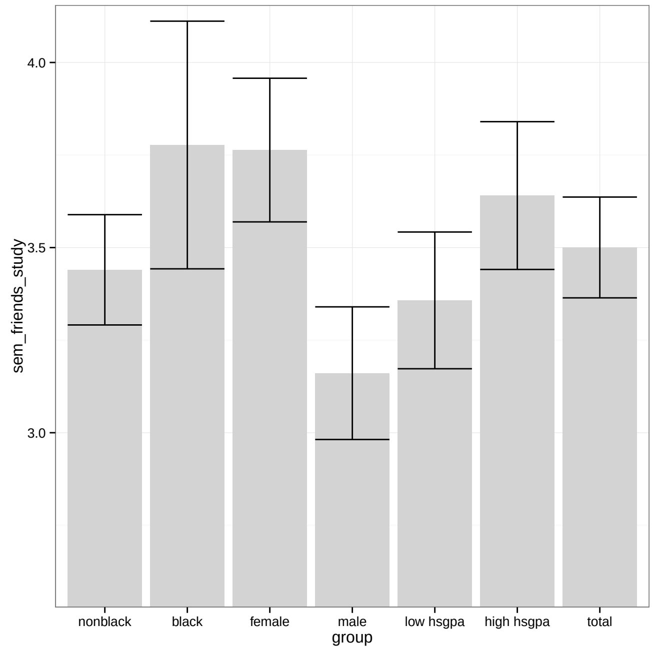


Table 3: Network characteristics

Churn	
Prob. friendship reported first semester but not second i.e. $\Pr\{A_2(i,j)=0 \mid A_1(i,j)=1\}$	0.51
Prob. second semester friendship is new i.e. $\Pr\{A_1(i,j)=0 \mid A_2(i,j)=1\}$	0.51

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

GPA (right column) given a set of individual and friend average variables. The study time regression shows evidence of significant partial correlations of own study semester-averaged study time with own gender, own high school GPA, own high school study time, and the average of friends' contemporaneous semester study time. Key to our analysis, own study time has a positive partial correlation with the average study time of one's friends. The GPA regression shows that own achievement has a significant positive partial correlation with being female, non-black, and having higher HS GPA. Own achievement has a significant partial correlation with own study time and the average of friends' contemporaneous semester study time. The results in Table 4 motivate our theoretical model. The regression of own study time is a rough approximation of student best response functions, where friend study time may affect own study time, even taking into account one's own characteristics. The regression of achievement on own and friend study times approximates an achievement production function, and suggests that both own and friend inputs may matter for achievement.

Appendix A.3 shows that interactions between own and friend characteristics do not affect the findings from Table 4. It also shows that some first semester inputs may enter second semester production, which we will deal with if/when we go dynamic. For now, note that contemporaneous (second semester) inputs still matter.

Table 4: Study and GPA OLS regressions, pooled over both semesters

	Dependent variable:	
	Own study	GPA
	(1)	(2)
Male	-0.336^{**}	-0.122^{**}
	(0.133)	(0.062)
Black	0.220	-0.389^{***}
	(0.172)	(0.079)
HS GPA	0.352^{**}	0.512^{***}
	(0.144)	(0.067)
HS Study	0.042^{***}	-0.001
	(0.006)	(0.003)
Own study		0.079^{***}
		(0.019)
Friend study	0.167^{***}	0.039^{**}
	(0.037)	(0.018)
Constant	1.292^{**}	0.907^{***}
	(0.537)	(0.249)
Observations	574	571
R^2	0.166	0.233
Adjusted R^2	0.158	0.225
Residual Std. Error	1.497 (df = 568)	0.686 (df = 564)
F Statistic	22.530^{***} (df = 5; 568)	28.561^{***} (df = 6; 564)
Note:		*p<0.1; **p<0.05; ***p<0.01

3 Model

Students are indexed by $i = 1, \ldots, I$ and time periods (semesters) by $t = 1, 2$. Denote observed variables with $\cdot^o$. Observed human capital created for student i in period t:

$$
\begin{aligned}
y_{it}^o =& y(s_{it}, s_{-it}, \mu_{yi}) + \epsilon_{it} \\
=& \beta_1 + \beta_2 s_{it} + \beta_3 s_{-it} + \mu_{yi} + \epsilon_{it}
\end{aligned} \tag{1}
$$

where human capital created that period, $y(s_{it}, s_{-it}, \mu_{yi})$, depends on student i's study time during period t, s_{it}; the average study time of i's peers, s_{-it}; and i's human capital type, $\mu_{yi} = x_i'\omega_y$, where x_i is i's vector of race, sex, high school GPA, combined ACT score, and high school study time.[4] The shock $\epsilon_{it} \sim \mathrm{N}\left(0, \sigma_\epsilon^2\right)$ is an iid measurement error for human capital created that period. Note that model, not observed, study times enter the technology. We include the direct effect of peer studying, s_{-it}, to make our specification comparable with other studies of peer effects.

Let A_t denote the adjacency matrix in period t. If the vector of all students' study times during that period is S_t, then average study time of i's peers during that period is: $s_{-it} \equiv 1_{it} A_t S_t / 1_{it} A_t$ where 1_{it} is a row vector of zeros with a 1 in the ith position.

Within-period timing:

1. Adjacency matrix A_t is determined
2. Students simultaneously choose own study time
3. Human capital is produced
4. Measurement errors for study time and human capital are realized

[4] We model friend study effort as the average of friend study time based on patterns we observed in the data. All the results about existence and uniqueness of equilibrium would still hold if *total* peer study time entered the student cost function.

A student chooses a series of study times (one per period) to maximize her objective

$$\{s_{it}^*\}_{t\in T} = \underset{\{s_t\}_{t\in T}}{\arg\max} \left\{ \sum_{t=1}^{2} y(s_t, s_{-it}, \mu_{yi}) - c(s_t, s_{-it}, \mu_{si}) \right\},$$

where $\mu_{si} = x_i'\omega_s$ is i's study type and effort cost $c(\cdot,\cdot,\cdot)$ is determined by

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

where $\gamma(\mu_{si}) = \exp(\tau_{\mu,1}\mu_{si} + \tau_{\mu,2}\mu_{si}^2)$ allows student policy functions to have different intercepts and slopes, based on their study type, and $\tau_s \leq 1$. The cost function $c(\cdot,\cdot,\cdot)$ is strictly convex in s_{it} and concave in s_{-it}. Peer study time enters the cost function by reducing the cost of one's own studying. It's less arduous (or more fun) to study at the library when everyone else is at the library. Conversely, if all your friends are at the library, it is no fun to stay in your dormitory room.

3.1 Model Solution

The student's decision problem is additively separate across time periods t. Therefore the student can solve each period's problem separately.[5] Student i's policy in t is

$$s_{it}^* = \arg\max_s \{y(s, s_{-it}, \mu_{yi}) - c(s, s_{-it}, \mu_{si})\} \tag{2}$$

which we can find by taking the first order condition with respect to own effort. Differentiating (2) in period t with respect to s gives $\frac{\partial y}{\partial s} = \frac{\partial c}{\partial s}$, i.e. the utility-maximizing study effort level equates the marginal return for increasing study effort with the marginal cost.

[5] If utility were nonlinear in semester achievement or the argument of the cost function were study time over the whole semester, the problem would no longer be separable across time periods. We assume student utility is linear in achievement because nonlinearity of utility in achievement would be difficult to separate from nonlinearity in the cost function, outside functional form restrictions.

Expanding the optimality condition gives:

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

$$\Rightarrow$$

$$s_{it}^* = -\theta_3 - \theta_4 \frac{1}{\gamma(\mu_{si})} + (\beta_2 - \theta_1) s_{-it}^{\tau_s} - \theta_2 \frac{s_{-it}^{\tau_s}}{\gamma(\mu_{si})} \tag{3}$$

which can be equivalently expressed in terms of the policy function:

$$s_{it}^* = \underbrace{\kappa_1}_{-\theta_3} + \underbrace{\kappa_2}_{-\theta_4} \frac{1}{\gamma(\mu_{si})} + \underbrace{\kappa_3}_{\beta_2 - \theta_1} s_{-it}^{\tau_s} + \underbrace{\kappa_4}_{-\theta_2} \frac{s_{-it}^{\tau_s}}{\gamma(\mu_{si})} \equiv \psi_i(s_{-it}) \tag{4}$$

where $(\kappa_1, \kappa_2, \kappa_3, \kappa_4)$ are composites of the structural technology parameter β_2 and cost function parameters $(\theta_1, \theta_2, \theta_3, \theta_4)$. We assume κ_1 and κ_3 are positive. Note the policy function is indexed by student because students have different best response functions $\psi(\cdot)$ based on their study type μ_s. Own effort is increasing linearly in the productivity of own effort β_2, and may also increase in own study type, μ_{si}, depending on κ_2 and κ_4. Own effort is an increasing and concave function of friend's effort.

3.1.1 Equilibrium

Definition 1 (Period Nash equilibrium)**.** *A pure strategy Nash equilibrium in study times* $S_t^* = [s_{1t}^*, s_{2t}^*, \cdots, s_{It}^*]'$ *satisfies* $s_{it}^* = \psi_i(s_{-it}^*)$*, for* $i \in I$*, given adjacency matrix* A_t*.*

Claim 1. *Let k be the number of hours during the entire time period t. There exists a unique pure strategy Nash equilibrium if $\psi_i : R^I \mapsto R$ are strictly concave and weakly increasing, $\psi_i(0) > 0$, and $\psi_i(k) < k$ for $i \in I$.*

Proof. Define $S = [0, k]^I$, i.e. a compact and convex set. Define a function Ψ:

$$\Psi : S \mapsto S = \begin{bmatrix} \psi_1(x_{-1}) - x_1 \\ \psi_2(x_{-2}) - x_2 \\ \vdots \\ \psi_I(x_{-I}) - x_I \end{bmatrix}$$

.

Existence: By Brouwer's Fixed Point Theorem, Ψ is a continuous self map on the compact set S, so an equilibrium exists.

Uniqueness: Note that Ψ is also strictly concave and because $\psi_i(k) < k$, we have $\Psi(K) < 0$, where K is an vector of length I whose elements are all k.

Say S^* solves $\Psi(S^*) = 0$. Take $\tilde{S} < S^*$ (i.e. no element larger, and at least one element strictly less). Due to the concavity of Ψ we also have that Ψ is quasiconcave, which means it has convex upper contour sets. This, combined with $\Psi(0) > 0$, implies that $\Psi(\tilde{S}) > 0$. Analogously, $\tilde{\tilde{S}} > S^*$ can't solve Ψ.

(This is intuitive in the two player case. Draw a picture.)

□

4 Likelihood

The endogenous objects are study times and GPAs. Recall that the adjacency matrix is predetermined, but may vary between periods. Also recall that all shocks are assumed to be independently distributed, though we stress that our analysis exploits an unusually rich set of covariates thought to affect both sorting into friendships and the evolution of study time and human capital, such as how much students studied during high school.

The model provides a mapping from the period adjacency matrix A_t to a unique equilibrium in study times, S_t. Given an equilibrium vector S_t^*, achievement in equilibrium y_{it}^* is computed by applying the technology $y(s_i, s_{-i}, \mu_y)$. Semester achievement is noisily measured by $y_{it}^o \in [0, 4]$ according to: $y_{it}^o = \max\{0, \min\{y_{it}^* + \epsilon_{it}, 4\}\}$, where $\epsilon \sim iid\, \mathrm{N}\,(0, \sigma_\epsilon^2)$ is classical measurement error. Each model period t has several study time surveys, indexed by $\tilde{t}$. Observed study time $s^o_{i\tilde{t}} \in [0, \infty)$ is noisily measured according to $s^o_{i\tilde{t}} = \max\{0, s^*_{it} + \eta_{i\tilde{t}}\}$, where the classical measurement error η is Gaussian.

The likelihood for GPA in semester t takes into account censoring at both 0 and 4:

$$L^y_{it} = \Phi\left(\frac{0 - y^*_{it}}{\sigma_\epsilon}\right)^{\mathbf{1}\{y^o_{it}=0\}} \times \left(1 - \Phi\left(\frac{4 - y^*_{it}}{\sigma_\epsilon}\right)\right)^{\mathbf{1}\{y^o_{it}=4\}} \times \frac{1}{\sigma_\epsilon}\phi\left(\frac{y^o_{it} - y^*_{it}}{\sigma_\epsilon}\right).$$

The likelihood for study time in survey $\tilde{t}$ (which took place in period t) takes into account censoring at 0:

$$L^s_{i\tilde{t}} = F_\eta\left(0 - s^*_{it}\right)^{\mathbf{1}\{s^o_{i\tilde{t}}=0\}} \times f_\eta\left(s^o_{i\tilde{t}} - s^*_{it}\right),$$

where f_η and F_η are the normal distribution density and CDF, respectively.

The likelihood for student i is therefore:

$$L_i = \left(\Pi_{\tilde{t}\in t} L^s_{i\tilde{t}}\right) \times \left(\Pi_t L^y_{it}\right).$$

Note that because there is no unobservable heterogeneity in the model, we need not integrate out student ability when calculating the likelihood L_i.

5 Estimation Results

Table 5 contains parameter estimates. The key technology parameter is $\beta_2 = 0.210$, the marginal product of own study effort on achievement. Put another way, increasing own (model) study time by one sd increases achievement by 39% sd (ceteris paribus).[6] Human capital type loadings, which are essentially regression coefficients for the associated variables, all have the same sign as the reduced-form approximation of the technology (Table 4): human capital type is increasing in high school GPA and ACT score, and Black students have lower

[6]We also estimated a specification of the model where friend study time directly entered the technology. We exclude it here because we found the coefficient to be close to zero.

levels of human capital growth.[7]

The study cost function coefficients $(\theta_1, \theta_2, \theta_3)$ are precisely estimated, but it is perhaps more intuitive to recast them in terms of the study time policy function (**??**). The intercept $\kappa_1 = 0.863$, satisfying the condition that $\psi_i(0) > 0, \forall i$. The study-type-specific intercept, κ_2, is governed by θ_4, which we estimated as close to zero and therefore excluded in the cost function. Own study time is increasing in friend study time, $\kappa_3 = 1.484$, and higher study types have steeper policy functions ($\kappa_4 = -0.976$). The study policy function is almost linear in friend study time $\tau_s = 0.988$. Being a higher study type increases the slope of the policy function ($\tau_{\mu,1} = 0.115$), but at a decreasing rate ($\tau_{\mu,2} = -0.005$). Study type is increasing in high school GPA and study time, and decreasing in combined ACT score. Though Black students study considerably more than non-Black students, the coefficient on being Black is negative. Males study less than females, even after taking high school characteristics into account.

To get intuition for heterogeneity in student studytime policy functions, Figure 3 plots policy functions for the 25th (red line) and 75th (blue line) percentile study types. Because $\kappa_1 > 0$ they both have positive intercepts, and because $\kappa_2 = 0$, both study types share the same intercept. However, because $\kappa_4 < 0$, the higher study type has a steeper policy function, that is, she is more responsive to an increase in friend study effort than the lower study type. By comparing where the low and high types' policy functions intersect the identity function (dotted black line), we can identify equilibria when each student is matched with her respective study type. When two 75th percentile study types are paired they would study almost 5 hours each, almost twice the amount two 25th percentile study types would study when paired together. The estimated study time policy functions indicate that the game is supermodular. Students study more, and therefore produce more human capital in total, when matched by type. That being said, whether or not students will take advantage of this supermodularity depends on how they sort into friendships.

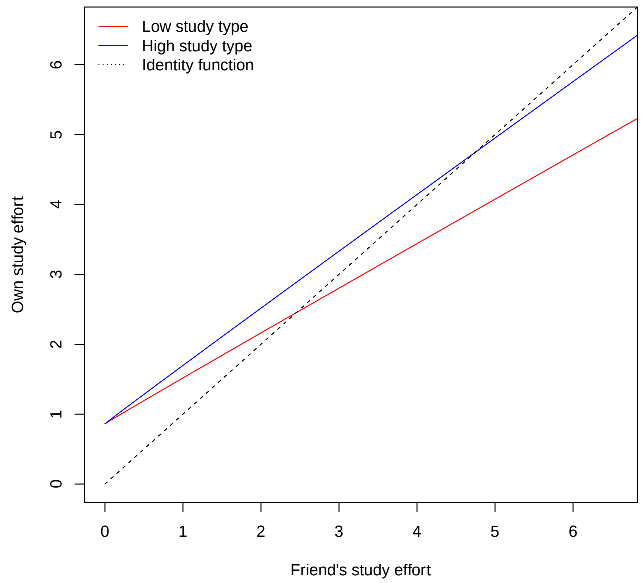
Figure 4 shows that the model closely fits observed study time and GPA, both overall

[7]Because high school study time and sex did not appreciably change the human capital type, the associated $\omega_{y\cdot}$ were set to zero.

Table 5: Parameter estimates

Parameter	Estimate	SE	Description
Technology			
β_1	-0.489	0.325	intercept
β_2	0.210	0.044	marginal product of own study time
$\omega_{y,\text{HS GPA}}$	0.450	0.073	loading of HS GPA on human capital type
$\omega_{y,\text{ACT}}$	0.053	0.01	loading of ACT on human capital type
$\omega_{y,\text{Black}}$	-0.244	0.089	loading of Black on human capital type
Study cost function			
θ_1	-1.275	0.114	study cost terms
θ_2	0.976	0.118	study cost terms
θ_3	-0.863	0.312	study cost terms
τ_s	0.988	0.063	curvature on friend study time
$\tau_{\mu,1}$	0.115	0.044	linear term for study type
$\tau_{\mu,2}$	-0.005	0.003	quadratic term for study type
$\omega_{s,\text{HS GPA}}$	1.000	—	loading of HS GPA on study type, fixed to 1
$\omega_{s,\text{ACT}}$	-0.099	0.042	loading of ACT on study type
$\omega_{s,\text{Black}}$	-0.612	0.332	loading of Black on study type
$\omega_{s,\text{Male}}$	-0.876	0.320	loading of Male on study type
$\omega_{s,\text{HS study}}$	0.254	0.059	loading of HS study on study type
Shocks			
σ_ϵ	0.732	0.018	sd measurement error for GPA
σ_η	2.164	0.025	sd measurement error for observed study time

Figure 3: Study policy function for 25th and 75th percentile study type μ_s



and by student subgroup.

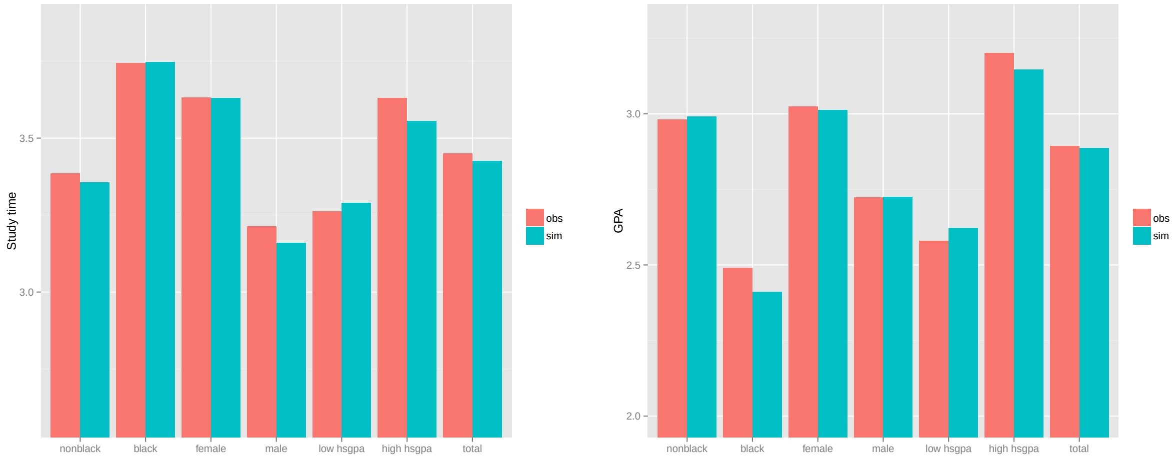
Table 6 shows that the model captures the relationship between own and friend study time, though it understates the relationship.

Table 6: Regression of own study time on student characteristics

	Own study time	
	Observed	Simulated
Black	0.128	0.174
Male	-0.227	-0.316
HS GPA	0.227	0.150
HS study time	0.035	0.041
Avg. friend study time	0.464	0.279
Constant	0.739	1.592

Note: Both own and friend study time are averaged over both semesters

Figure 4: Fit of study time (left) and GPA (right)



6 Quantitative Findings

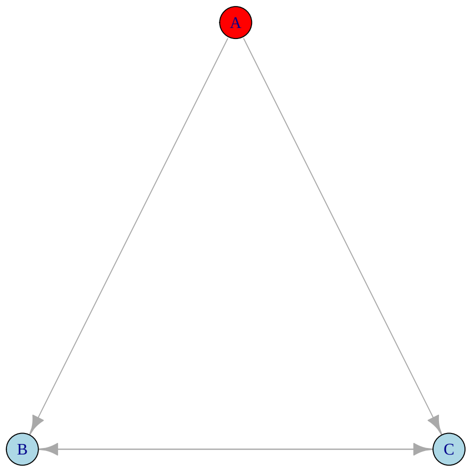
We conduct two simulations to quantify the role of social interactions in the production of achievement in our context. First, we investigate how general equilibrium effects augment social interactions, by exogenously increasing the study time of a student by one standard deviation and measuring how study times and achievement change for connected peers, with and without taking into account general equilibrium effects. Second, because peer effects manifest through 1) sorting into friendships and 2) the effects of inputs chosen by friends, we compare achievement in the baseline model with that when friends are instead unconditionally randomly assigned.

6.1 Impulse response function

Partial versus general equilibrium effects

Differentiate y_B with respect to s_A:

$$\frac{\partial y_B}{\partial s_A} = \beta_2 \frac{\partial s_B}{\partial s_{-B}} \frac{\partial s_{-B}}{\partial s_A} = \beta_2 \frac{\partial s_B}{\partial s_{-B}} \frac{\partial (\frac{s_A}{2} + \frac{s_C}{2})}{\partial s_A} = \beta_2 \psi_B'(s_{-B}) \left(\frac{1}{2} + \frac{\psi_C'(s_{-C})}{2} \right)$$



"No equilibrium effect" (other than A on B):

- We can use the estimated policy function and technology to calculate how y_B changes due to a change in s_A, fixing $s_C = s_C^0$
- If $\psi'_C(s_{-C}) > 0$, this will understate the effect of increasing s_A
- How important is this equilibrium effect?

Figure 5 demonstrates the difference between partial and general equilibrium effects. First, consider the initial equilibrium characterized by the intersection of best responses of B and C, given $s_A = s_A^0$ (black dot). The increase in s_A shifts ψ_B vertically for every s_C. The partial equilibrium effect ignores the change in ψ_C (red circle). The general equilibrium effect takes into account change in ψ_C (blue dot), and is larger.

Though in theory, general equilibrium effects will be larger than partial equilibrium effects because study time is increasing for all students in friend study time, the extent to which partial and general equilibrium effects differ is an empirical question. To quantitatively examine the difference, we consider an exogenous increase of one student's study time. Table 7 shows the results, expressed in standard deviations of the outcome variable.

Figure 5: Partial versus general equilibrium effects

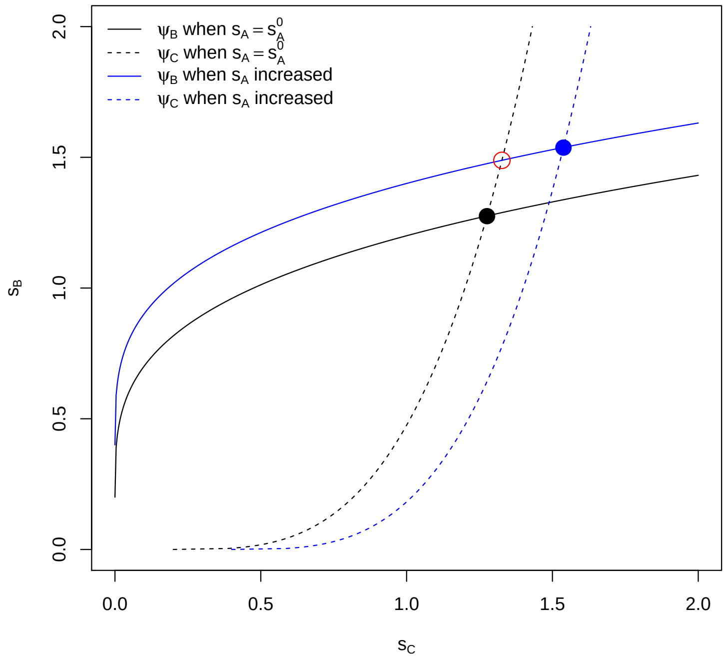


Table 7: Partial versus general equilibrium effects of an exogenous increase in one student's study time (%sd)

	GPA		Own study time		Friend study time	
Dist. from shocked node	$= 1$	≤ 2	$= 1$	≤ 2	$= 1$	≤ 2
First iteration	0.16	0.10	0.28	0.14	0.44	0.30
New equilibrium	0.18	0.14	0.34	0.23	0.54	0.37
Equil. gain	22%	36%	22%	67%	22%	23%

- Total effect bigger for students directly connected to shocked student (shock diluted for others)

- Effect on friend study time is larger than effect on own

- Gain from using equilibrium bigger for students not directly connected

6.2 Randomly assigning friends

Table 8 shows that random assignment works.

Table 8: Characteristics of social network

	Baseline	Random friends
Avg. number of friends	3.09	3.42
Correlations bt. own and avg. friends		
Black	0.74	0.04
Male	0.71	0.09
HS GPA	0.23	-0.04
ACT	0.31	0.06
HS study	0.23	0.04

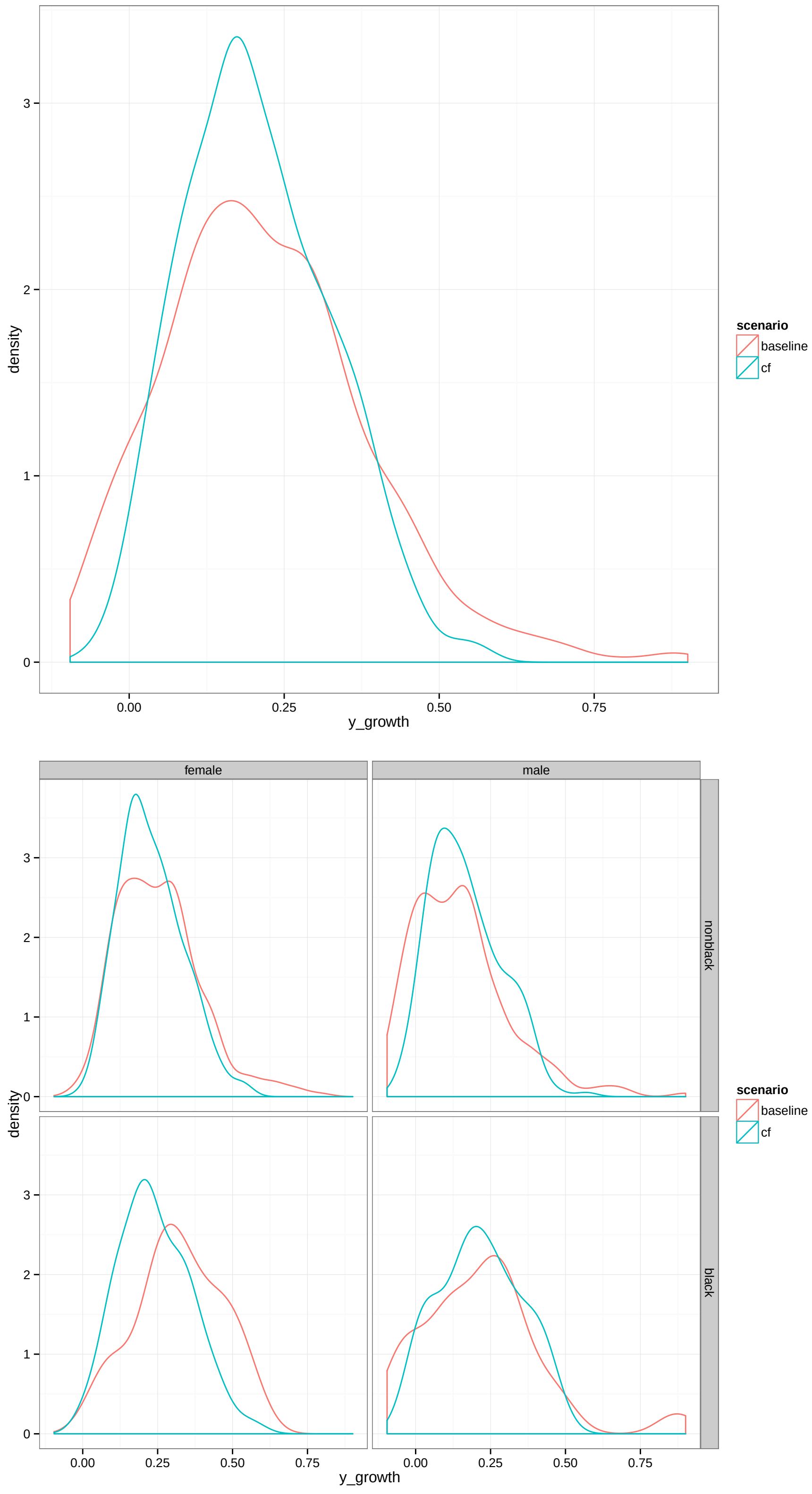
Table 9 summarizes the changes in model outcomes, expressed in standard deviations of the outcome variable in the baseline.

Table 9: Change in outcomes from randomly assigning friends, in sd of outcome variable

	Achievement	Achievement, net μ_y	Own study time	Friend study time
Total	-0.03	-0.08	-0.09	-0.10
% change sd	-0.04	-0.28	-0.28	-0.44
Nonblack	-0.01	-0.04	-0.06	-0.02
Black	-0.11	-0.28	-0.24	-0.45
Female	-0.08	-0.21	-0.21	-0.32
Male	0.03	0.08	0.07	0.19
Below med. HS GPA	0.01	0.03	0.02	0.09
Above med. HS GPA	-0.07	-0.19	-0.18	-0.28

Figure 6 plots the density of achievement production, net human capital type μ_y, to examine how sorting into friendships affects the production of human capital. The red curve is the density of achievement growth in the baseline and the blue is achievement growth under the random friends counterfactual. The variance of achievement is smaller under the counterfactual for two reasons: there is no longer sorting into friendships along study type, and the supermodularity fo the studytime game further intensifies this effect. The lower panel plots both densities, splitting students by sex and race. We can clearly see that black females gain the most from non-random sorting into friendships.

Figure 6: Distribution of achievement, net human capital type μ_y, (y_growth) under baseline and random friends counterfactual



A Data

A.1 Survey questions

A.2 Other data details

- Currently we use only one friend survey per semester. FS3 is the first semester friends, FS4 is the second semester friends.
  - We will later check robustness using other FS and assumptions about the relevance of the different FS.
- To most closely match the way friends enter in the model, links have been symmetrized.
- By setting the model period to a semester, we effectively say that reporting noise causes variation within a semester in reported study time.
  - Regress $s^o_{i\tau} = \beta_0 + \beta_1 \overline{s^o}_{it} + \sum_\tau \beta_k \mathbf{1}\{k = \tau\} + \epsilon_{i\tau}$, where survey τ is in semester t, so β_k is a dummy variable for being in survey k, and $\overline{s^o}_{it}$ is mean reported study time for that student in semester t.
    - The standard deviation of the residual from the above regression is 1.47. The standard deviation of $\overline{s}_{it}$ between students and semesters is 1.67. Therefore, between-student variance in reported study time is 29% (=$1 - (\frac{1.67}{1.47})^2$) higher than within-student variance.

A.3 Testing for interaction effects in study time, GPA regressions

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

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

(6

Table 10 estimates the study time regression in Table 4, allowing for interactions between own characteristics and friend study time (columns 2-4). Other than males, who are less responsive to increases in friend study time than other students, there are no significant differences between groups.

Table 10: Pooled over both semesters, interactions

	Dependent variable:			
	sem_study			
	(1)	(2)	(3)	(4)
male	−0.154	−0.153	0.433	−0.154
	(0.137)	(0.137)	(0.293)	(0.137)
black	0.414**	0.553	0.411**	0.414**
	(0.177)	(0.394)	(0.177)	(0.178)
hsgpa	0.325**	0.330**	0.334**	0.328
	(0.147)	(0.148)	(0.147)	(0.294)
sem_friends_sem_study	0.181***	0.189***	0.267***	0.184
	(0.044)	(0.048)	(0.058)	(0.292)
I(black *sem_friends_sem_study)		−0.043		
		(0.109)		
I(male *sem_friends_sem_study)			−0.198**	
			(0.087)	
I(hsgpa *sem_friends_sem_study)				−0.001
				(0.085)
Constant	1.835***	1.794***	1.523***	1.826*
	(0.542)	(0.552)	(0.558)	(1.018)
Observations	646	646	646	646
R^2	0.052	0.052	0.059	0.052
Adjusted R^2	0.046	0.044	0.052	0.044
Residual Std. Error	1.627 (df = 641)	1.628 (df = 640)	1.622 (df = 640)	1.629 (df = 640)
F Statistic	8.726*** (df = 4; 641)	7.003*** (df = 5; 640)	8.053*** (df = 5; 640)	6.970*** (df = 5; 640)
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 14 does the same as Table 10, but for GPA. None of the estimated interactions of one's own characteristics on either own or friend study time are discernible from zero at the 5% level.

Finally, Table 15 augments the regressions in Table 4 by adding in mean friend characteristics. The takeaway is that friend study time, not friend characteristics, is what correlates with changes in own study time (column 2) and achievement (column 4).

Table 16 shows that first semester inputs may enter the production of GPA in the second semester.

Table 17 shows that a measure of study productivity, high school GPA divided by high school study time, does not substantially affect either the relationship between friend and own study time (column 1) or own and friend study time and the production of GPA (column

Table 11: Study regressions interacted with own characteristics, pooled over both semesters

	Dependent variable:			
	sem_study			
	(1)	(2)	(3)	(4)
male	-0.340^{**}	-0.072	-0.337^{**}	-0.342^{**}
	(0.134)	(0.303)	(0.134)	(0.134)
black	0.560	0.217	0.220	0.218
	(0.399)	(0.172)	(0.172)	(0.172)
hsgpa	0.361^{**}	0.360^{**}	0.326	0.352^{**}
	(0.145)	(0.144)	(0.297)	(0.144)
studyhs	0.042^{***}	0.042^{***}	0.042^{***}	0.035^{***}
	(0.006)	(0.006)	(0.006)	(0.012)
sem_friends_study	0.183^{***}	0.192^{***}	0.143	0.139^{**}
	(0.041)	(0.045)	(0.251)	(0.057)
I(black ∗sem_friends_study)	-0.088			
	(0.094)			
I(male ∗sem_friends_study)		-0.077		
		(0.079)		
I(hsgpa ∗sem_friends_study)			0.007	
			(0.072)	
I(studyhs ∗sem_friends_study)				0.002
				(0.003)
Constant	1.208^{**}	1.167^{**}	1.380	1.392^{**}
	(0.544)	(0.552)	(1.040)	(0.558)
Observations	574	574	574	574
R^2	0.167	0.167	0.166	0.166
Adjusted R^2	0.158	0.158	0.157	0.157
Residual Std. Error (df = 567)	1.497	1.497	1.498	1.498
F Statistic (df = 6; 567)	18.919^{***}	18.932^{***}	18.744^{***}	18.828^{***}
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 12: Study regressions interacted with friend characteristics, pooled over both semesters

	Dependent variable:			
	sem_study			
	(1)	(2)	(3)	(4)
male	-0.359^{***}	-0.346^{**}	-0.343^{**}	-0.327^{**}
	(0.134)	(0.134)	(0.133)	(0.133)
black	0.565^{**}	0.170	0.220	0.247
	(0.277)	(0.184)	(0.172)	(0.173)
hsgpa	0.333^{**}	0.337^{**}	0.349^{**}	0.367^{**}
	(0.145)	(0.146)	(0.144)	(0.144)
studyhs	0.041^{***}	0.041^{***}	0.034^{***}	0.047^{***}
	(0.006)	(0.006)	(0.007)	(0.007)
sem_friends_study	0.183^{***}	0.161^{***}	0.137^{***}	0.173^{***}
	(0.039)	(0.038)	(0.042)	(0.038)
I((black ∗sem_friends_black) ∗sem_friends_study)	-0.130			
	(0.082)			
I((black - sem_friends_black)ˆ2 ∗sem_friends_study)		0.070		
		(0.090)		
I((studyhs ∗sem_friends_studyhs) ∗sem_friends_study)			0.0001	
			(0.0001)	
I((studyhs - sem_friends_studyhs)ˆ2 ∗sem_friends_study)				-0.0001
				(0.0001)
Constant	1.317^{**}	1.370^{**}	1.428^{***}	1.191^{**}
	(0.536)	(0.546)	(0.543)	(0.541)
Observations	574	574	574	574
R^2	0.169	0.166	0.169	0.168
Adjusted R^2	0.160	0.158	0.161	0.159
Residual Std. Error (df = 567)	1.495	1.498	1.495	1.496
F Statistic (df = 6; 567)	19.247^{***}	18.860^{***}	19.259^{***}	19.121^{***}
Note:		*p<0.1; **p<0.05; ***p<0.01		

Table 13: Study regressions interacted with friend characteristics, pooled over both semesters

	Dependent variable:			
	sem_study			
	(1)	(2)	(3)	(4)
male	-0.359^{***}	-0.114	-0.328^{**}	-0.343^{**}
	(0.134)	(0.234)	(0.134)	(0.133)
black	0.565^{**}	0.226	0.241	0.220
	(0.277)	(0.172)	(0.174)	(0.172)
hsgpa	0.333^{**}	0.368^{**}	0.222	0.349^{**}
	(0.145)	(0.145)	(0.224)	(0.144)
studyhs	0.041^{***}	0.042^{***}	0.042^{***}	0.034^{***}
	(0.006)	(0.006)	(0.006)	(0.007)
sem_friends_study	0.183^{***}	0.185^{***}	0.046	0.137^{***}
	(0.039)	(0.040)	(0.164)	(0.042)
I((black ∗sem_friends_black) ∗sem_friends_study)	-0.130			
	(0.082)			
I((male ∗sem_friends_male) ∗sem_friends_study)		-0.088		
		(0.076)		
I((hsgpa ∗sem_friends_hsgpa) ∗sem_friends_study)			0.010	
			(0.013)	
I((studyhs ∗sem_friends_studyhs) ∗sem_friends_study)				0.0001
				(0.0001)
Constant	1.317^{**}	1.166^{**}	1.741^{**}	1.428^{***}
	(0.536)	(0.548)	(0.798)	(0.543)
Observations	574	574	574	574
R^2	0.169	0.167	0.166	0.169
Adjusted R^2	0.160	0.159	0.158	0.161
Residual Std. Error (df = 567)	1.495	1.497	1.498	1.495
F Statistic (df = 6; 567)	19.247^{***}	19.011^{***}	18.857^{***}	19.259^{***}
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 14: Pooled over both semesters, interactions

	Dependent variable:			
	gpa			
	(1)	(2)	(3)	(4)
male	-0.131^{**}	-0.133^{**}	-0.052	-0.118^{*}
	(0.061)	(0.061)	(0.168)	(0.061)
black	-0.337^{***}	-0.570^{**}	-0.334^{***}	-0.328^{***}
	(0.079)	(0.244)	(0.079)	(0.079)
hsgpa	0.567^{***}	0.560^{***}	0.566^{***}	0.987^{***}
	(0.065)	(0.066)	(0.066)	(0.175)
sem_study	0.064^{***}	0.066^{***}	0.048^{*}	0.264^{*}
	(0.017)	(0.019)	(0.026)	(0.136)
I(black ∗sem_study)		-0.012		
		(0.050)		
I(male ∗sem_study)			0.026	
			(0.035)	
I(hsgpa ∗sem_study)				-0.058
				(0.039)
sem_friends_sem_study	0.054^{***}	0.039^{*}	0.081^{***}	0.297^{**}
	(0.020)	(0.022)	(0.027)	(0.133)
I(black ∗sem_friends_sem_study)		0.084^{*}		
		(0.051)		
I(male ∗sem_friends_sem_study)			-0.057	
			(0.040)	
I(hsgpa ∗sem_friends_sem_study)				-0.071^{*}
				(0.039)
Constant	0.693^{***}	0.756^{***}	0.664^{**}	-0.743
	(0.242)	(0.249)	(0.258)	(0.605)
Observations	636	636	636	636
R^2	0.228	0.232	0.231	0.237
Adjusted R^2	0.222	0.223	0.223	0.228
Residual Std. Error	0.714 (df = 630)	0.714 (df = 628)	0.714 (df = 628)	0.712 (df = 628)
F Statistic	37.299^{***} (df = 5; 630)	27.072^{***} (df = 7; 628)	26.981^{***} (df = 7; 628)	27.849^{***} (df = 7; 628)
Note:				$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 15: Pooled over both semesters

	Dependent variable:			
	sem_study		gpa	
	(1)	(2)	(3)	(4)
male	−0.154	0.002	−0.131**	−0.178**
	(0.137)	(0.198)	(0.061)	(0.088)
black	0.414**	0.357	−0.337***	−0.253**
	(0.177)	(0.254)	(0.079)	(0.115)
hsgpa	0.325**	0.335**	0.567***	0.521***
	(0.147)	(0.148)	(0.065)	(0.067)
sem_study			0.064***	0.076***
			(0.017)	(0.018)
sem_friends_sem_study	0.181***	0.208***	0.054***	0.049**
	(0.044)	(0.047)	(0.020)	(0.021)
sem_friends_black		0.225		−0.105
		(0.319)		(0.143)
sem_friends_male		−0.283		0.080
		(0.258)		(0.115)
sem_friends_hsgpa		0.332		0.101
		(0.238)		(0.107)
Constant	1.835***	0.601	0.693***	0.473
	(0.542)	(0.958)	(0.242)	(0.429)
Observations	646	619	636	614
R^2	0.052	0.071	0.228	0.229
Adjusted R^2	0.046	0.061	0.222	0.219
Residual Std. Error	1.627 (df = 641)	1.584 (df = 611)	0.714 (df = 630)	0.706 (df = 605
F Statistic	8.726*** (df = 4; 641)	6.688*** (df = 7; 611)	37.299*** (df = 5; 630)	22.429*** (df = 8; 6
Note:				*p<0.1; **p<0.05; ***p<

Table 16: Evidence of human capital dynamics

	Dependent variable:	
	gpa2	
	(1)	(2)
black	-0.275^{**}	-0.039
	(0.115)	(0.103)
male	-0.078	0.009
	(0.091)	(0.079)
hsgpa	0.544^{***}	0.282^{***}
	(0.096)	(0.088)
studyhs	-0.003	-0.0001
	(0.004)	(0.003)
sem2_study	0.096^{***}	0.058^{**}
	(0.031)	(0.025)
sem2_frstudy	0.041^{**}	0.031^{*}
	(0.020)	(0.018)
sem1_study	0.045	
	(0.031)	
gpa1		0.562^{***}
		(0.059)
Constant	0.568	0.021
	(0.362)	(0.318)
Observations	271	275
R^2	0.265	0.455
Adjusted R^2	0.246	0.441
Residual Std. Error	0.680 (df = 263)	0.595 (df = 267)
F Statistic	13.571^{***} (df = 7; 263)	31.834^{***} (df = 7; 267)
Note:		*p<0.1; **p<0.05; ***p<0.01

2).

Table 18 shows that excluding friend study time does not appreciably change the relationship between own study time and GPA, even controlling for friend characteristics.

Table 17: Study and GPA regressions, pooled over both semesters

	Dependent variable:	
	sem_study	gpa
	(1)	(2)
male	-0.307^{**}	-0.081
	(0.138)	(0.064)
black	0.210	-0.395^{***}
	(0.174)	(0.080)
hsgpa	0.353^{**}	0.498^{***}
	(0.148)	(0.069)
studyhs	0.038^{***}	-0.003
	(0.006)	(0.003)
sem_study		0.077^{***}
		(0.022)
sem_friends_study	0.142^{***}	0.042^{**}
	(0.045)	(0.021)
I(hsgpa/studyhs)	-0.186	0.010
	(0.132)	(0.067)
I(hsgpa ∗sem_study/studyhs)		-0.005
		(0.014)
I(hsgpa ∗sem_friends_study/studyhs)	0.040	-0.006
	(0.038)	(0.018)
Constant	1.474^{***}	0.990^{***}
	(0.550)	(0.256)
Observations	546	543
R^2	0.162	0.221
Adjusted R^2	0.151	0.208
Residual Std. Error	1.494 (df = 538)	0.684 (df = 533)
F Statistic	14.857^{***} (df = 7; 538)	16.822^{***} (df = 9; 533)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 18: Study and GPA regressions with only friend characteristics, pooled over both semesters

	Dependent variable:			
	sem_study	gpa	sem_study	gpa
	(1)	(2)	(3)	(4)
male	−0.304	−0.138	−0.324*	−0.141
	(0.186)	(0.087)	(0.187)	(0.087)
black	0.210	−0.282**	0.195	−0.286**
	(0.252)	(0.118)	(0.254)	(0.118)
hsgpa	0.316**	0.496***	0.330**	0.498***
	(0.145)	(0.068)	(0.146)	(0.068)
studyhs	0.037***	−0.001	0.038***	−0.001
	(0.006)	(0.003)	(0.006)	(0.003)
sem_study		0.077***		0.083***
		(0.020)		(0.019)
sem_friends_study	0.127***	0.040**		
	(0.039)	(0.018)		
sem_friends_black	−0.056	−0.136	0.026	−0.113
	(0.310)	(0.145)	(0.311)	(0.145)
sem_friends_male	0.002	0.046	−0.065	0.025
	(0.235)	(0.109)	(0.236)	(0.109)
sem_friends_hsgpa	0.251	0.103	0.315	0.122
	(0.226)	(0.105)	(0.227)	(0.105)
sem_friends_studyhs	0.033***	0.0002	0.041***	0.002
	(0.009)	(0.004)	(0.009)	(0.004)
Constant	0.399	0.608	0.518	0.639
	(0.908)	(0.422)	(0.915)	(0.423)
Observations	574	571	574	571
R^2	0.187	0.236	0.172	0.230
Adjusted R^2	0.174	0.223	0.160	0.218
Residual Std. Error	1.483 (df = 564)	0.687 (df = 560)	1.495 (df = 565)	0.690 (df = 561
F Statistic	14.428*** (df = 9; 564)	17.337*** (df = 10; 560)	14.646*** (df = 8; 565)	18.620*** (df = 9;
Note:				*p<0.1; **p<0.05; ***p<

Table 19 shows that though production of achievement may look different across semesters, the effect of own study time on achievement becomes similar when we control for first semester GPA.

Table 19: Study and GPA regressions, by semester

	Dependent variable:					
	sem_study	gpa	sem_study	gpa	gpa2	
	(1)	(2)	(3)	(4)	(5)	(6)
male	-0.381^{**}	-0.164^{*}	-0.291	-0.072	-0.078	0.009
	(0.187)	(0.084)	(0.192)	(0.090)	(0.091)	(0.079)
black	0.220	-0.453^{***}	0.179	-0.308^{***}	-0.275^{**}	-0.039
	(0.246)	(0.111)	(0.243)	(0.115)	(0.115)	(0.103)
hsgpa	0.323	0.477^{***}	0.341^{*}	0.563^{***}	0.544^{***}	0.282^{***}
	(0.208)	(0.094)	(0.203)	(0.096)	(0.096)	(0.088)
studyhs	0.040^{***}	0.0004	0.041^{***}	-0.002	-0.003	-0.0001
	(0.008)	(0.004)	(0.008)	(0.004)	(0.004)	(0.003)
sem_study		0.045^{*}		0.121^{***}		
		(0.026)		(0.028)		
sem_friends_study	0.269^{***}	0.013	0.141^{***}	0.046^{**}		
	(0.081)	(0.037)	(0.042)	(0.020)		
sem2_study					0.096^{***}	0.058^{**}
					(0.031)	(0.025)
sem2_frstudy					0.041^{**}	0.031^{*}
					(0.020)	(0.018)
sem1_study					0.045	
					(0.031)	
gpa1						0.562^{***}
						(0.059)
Constant	1.110	1.243^{***}	1.396^{*}	0.540	0.568	0.021
	(0.764)	(0.344)	(0.760)	(0.363)	(0.362)	(0.318)
Observations	296	296	278	275	271	275
R^2	0.175	0.217	0.163	0.268	0.265	0.455
Adjusted R^2	0.161	0.201	0.148	0.252	0.246	0.441
Residual Std. Error	1.527 (df = 290)	0.684 (df = 289)	1.474 (df = 272)	0.688 (df = 268)	0.680 (df = 263)	0.595 (df = 267)
F Statistic	12.290^{***} (df = 5; 290)	13.368^{***} (df = 6; 289)	10.592^{***} (df = 5; 272)	16.352^{***} (df = 6; 268)	13.571^{***} (df = 7; 263)	31.834^{***} (df = 7; 267)

Note: *p<0.1; **p<0.05; ***p<0.01