NBER WORKING PAPER SERIES

PEER EFFECTS WITH RANDOM ASSIGNMENT: RESULTS FOR DARTMOUTH ROOMMATES

Bruce Sacerdote

Working Paper 7469
http://www.nber.org/papers/w7469

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2000

I would like to thank Phillip Hobbie and Mick Spencer at Dartmouth's Computing Services group and Lynn Rosenblum in Residential Life for helping me assemble the data. Thank you to Patty Anderson, Gary Engelhardt, Ed Glaeser, Larry Katz, Steven Levitt, and Doug Staiger for helpful comments and encouragement. Thank you to Hilla Talati and Michele Verni for their excellent assistance. This is a draft; suggestions are welcome and can be sent to bruce.sacerdote@dartmouth.edu. The views expressed herein are those of the author and not necessarily those of the National Bureau of Economic Research.

Peer Effects with Random Assignment:
Results for Dartmouth Roommates
Bruce Sacerdote
NBER Working Paper No. 7469
January 2000
JEL No.I2, J0

ABSTRACT

This paper uses a unique data set to measure peer effects among college age roommates. Freshman year roommates and dormmates are randomly assigned at Dartmouth College. I find that in this group, peer effects are very important in determining levels of academic effort and in decisions to join social groups such as fraternities. Residential peer effects are markedly absent in other major life decisions such as choice of college major. Several forms of peer effects are considered. The data support a model in which peer effects are driven by roommate behavior *after* the freshmen arrive. Social learning based on a roommate's *observable pre-Dartmouth* information or skills appears to be less important. Peer effects in GPA occur at the individual room level whereas peer effects in fraternity membership occur both at the room level and the entire dorm level. I also find that a freshman with high social ability is likely to remain with his or her roommates in sophomore year, but high academic ability actually decreases roommate retention.

Bruce Sacerdote
6106 Rockefeller Hall
Department of Economics
Dartmouth College
Hanover, NH 03755-3514
and NBER
Bruce.I.Sacerdote@dartmouth.edu

I. Introduction

Becker's work has revolutionized the social sciences by postulating that human behavior in a wide variety of areas can be understood as individual optimization subject to constraints. Noticeably absent from much of the work that has followed is a discussion of the importance of social interactions in determining individual behavior. Many models in the economics literature are based upon individual optimization without regard to what friends, neighbors, and other nearby actors in the economy are doing. This is most likely caused not by a belief that social interactions are unimportant, but more by the fact that it is difficult to model social interactions theoretically and to measure social interactions (peer effects) empirically.

This paper demonstrates and measures the importance of peer effects in a setting where peers are randomly assigned. Freshmen entering Dartmouth College are randomly assigned to dorms and to roommates. This eliminates the selection problem inherent in most data sets in which peers normally select each other based on observable and unobservable characteristics.

Furthermore by examining a range of outcomes, I am able to differentiate sharply between areas where peer effects are important for this group (eg level of academic effort, membership in social organizations) and areas that are unaffected by roommate and dormmate influences (eg choice of college major). While peer effects are large for outcomes such as joining fraternities, effects are smaller for outcomes that directly affect labor market activities; the effects on GPA are modest and there is no effect on choice of college major.

Following Manski (1993), I test whether the peer effects are driven by the roommate's *background* versus roommate's behavior *at Dartmouth* and I find in favor of the latter. I also find some evidence that students who do not express strong preferences (pre-treatment) regarding fraternities and amount of studying are more likely to be influenced by their roommates than students who arrive with strong preferences. Finally, the data suggest that students are more likely to remain with a roommate who provides positive social externalities as opposed to positive academic externalities.

Difficulties in measuring peer effects

The standard approach to measuring peer effects takes observational data and regresses own outcomes (or behaviors) on peer outcomes (or behaviors). For example Case and Katz (1991) regress criminal behavior, drug use, and church attendance on neighborhood averages for these variables. In another example, Kremer (1997) looks at the effects of parental and neighborhood educational attainment on youth educational attainment.

There are several difficulties inherent with this approach as detailed in Manski (1993). First, individuals generally self select into neighborhoods, groups, or roommate pairs. This makes it difficult to separate out the selection effect from any actual peer (treatment) effects. Secondly, if roommates i and j affect each others' GPAs simultaneously then it is difficult to separate out the actual causal effect that i has on j's outcome. Thirdly, note that correlation in outcomes may be driven by individuals' backgrounds (Manski calls these contextual effects) as opposed to events that occur during the observation period. The researcher may wish to distinguish between these two types of effects [2]

[2] For the discussion that follows I call the first issue "the selection problem" and the second issue "the endogeneity problem." The third issue is a matter of distinguishing between peer effects driven by *pre-treatment* characteristics and peer effects driven by events that occur *during treatment*.

Manski's language is slightly more technical. Manski recognizes three possible effects: a.) *endogenous effects* are driven by events that occur during treatment or observation. b.) *Contextual effects* are driven by the background of peers. c.) *Correlated effects* are driven by selection of individuals with similar backgrounds into a group. In my discussion, endogenous and contextual effects are two broad classes of peer effects. My "endogeneity problem" is what Manski calls the *reflection problem.*

Several authors attempt to solve the endogeneity problem by designing instruments for peer behavior which are assumed to be exogenous. For example Case and Katz (1991) and Gaviria and Raphael (1999) instrument for peer behavior using the average behavior of the peers' parents. Borjas (1995) regresses own behavior on measures of average human capital in the prior generation of one's ethnic group.[3]

Evans, Oates and Schwab attempt to solve the selection problem by adding an equation to explicitly model the fact that the teens in their data (a subsample of the NLSY) self-select into their peer group. While the aforementioned studies yield interesting and useful results, it is difficult to be certain about the exogeneity of the instruments or the ability of structural models to remove selection problems and deliver consistent estimates of peer effects.

This paper uses the random assignment of roommates to solve the selection problem inherent in most observational studies. Since roommates are randomly assigned, the selection problem is eliminated. And since I have data for earlier years in which there is selection (prior to the use of randomization) I can measure the importance of selection bias by comparing coefficients with and without selection bias. Random assignment implies that all of a roommates' background variables are uncorrelated with own background characteristics. This allows me to measure the causal effect of student i's background on his roommate j's outcomes.

I solve the endogeneity problem via a simple structural model. In the two roommate case, the model has a useful symmetry which implies restrictions on the variance covariance matrix of the residuals. This yields enough information to identify the effect of j's outcomes on i's outcomes thereby solving the endogeneity problem.

---

[3] In Manski's language, these author's are assuming no contextual effects in order to estimate the endogenous effects.

Different mechanisms for peer effects

The model allows me to identify separately the effects of *j's background* and j's *outcomes* on i's outcomes. This is Manski's distinction between contextual and endogenous effects. A peer effect based on background characteristics would likely involve a form of social learning as in Ellison and Fudenberg 1995, Banerjee 1992 or Griliches 1958. The idea here is that freshman arrive with heterogenous sets of knowledge about the world and about how to succeed at Dartmouth. For example, student i with excellent academic skills might transmit some of those skills to student j who arrives with a different set of skills.

Peer effects caused by during-treatment behavior (outcomes) could work through a variety of mechanisms such as information gathering, agglomeration externalities, or endogenous preference formation. Suppose i's information gathering at Dartmouth affects both i and j's outcomes as in Young (1993). Having roommates and dormmates explore various potential majors might generate information which would cause roommates together to switch into those fields where the signals were positive.[4]

A second possible source of during-treatment peer effects is agglomeration externality. In this model, when my roommate joins a fraternity, it raises the benefits to me of joining because I want to spend social time with my roommate in future years. Or it lowers the costs to me of joining since I already know one person in the organization. A final form of during-treatment peer effect that may be at work is endogenous preference formation as in Weber (1978), Romer (1999), and Glaeser (1999). This is a peer effect which works through roommates jointly determining their underlying preferences for hard work or joining fraternities. For a comprehensive discussion of these various forms of peer effects and related measurement issues see Glaeser and Scheinkman (1998).

Testing between the three during-treatment models is difficult and speculative at best. For example, I do find a strong effect in which student i is highly likely to join the same fraternity as his randomly assigned roommate j. And this effect does not work through j's observable background. However the effect could easily be driven by a.) agglomeration externality of joining the same House, b.) information that i or j gathers and shares about the specific House, or c.) a deeper shift in the preferences i and j both have which then makes joining that House more desirable.

Applications to Peer Effects More Broadly

It is important to ask to what degree the results in this paper can be generalized to other settings and there are certainly a number of caveats worth noting. The size and nature of peer effects in primary and secondary schooling are vital to thinking about what policy changes could be effective in improving outcomes in a given school. (See for example Betts and Morell 1998, Kain, Hanushek, and Rivkin 1998, Peterson 1997). The setting in this paper differs from a secondary school setting on at least three important dimensions. First, the students are older and hence perhaps less influenced by peer effects. Secondly the students live on campus rather than at home.

Finally, because of the highly selective admissions process, there is naturally less variation in academic ability among Dartmouth students than within a typical U.S. high school. It is not obvious whether this would increase or decrease the magnitude of peer effects. On the one hand, more variation leads to more possibilities for information to be exchanged. But, students may be less open to receiving information from a peer radically different from oneself.

---

[4] Good signals here are things like interesting material, fair grading, and good potential jobs upon graduation.

Empirical Framework

The data are analyzed using a basic model in which own GPA depends on own level of academic ability, roommate's level of ability, and roommate's GPA. The advantage of this approach is that it allows me to derive consistent estimates for the effects of roommate background and roommate GPA. (Subject to the structural assumptions of the model, this solves the endogeneity problem of regression i's GPA on j's GPA.[5]) I assume implicitly that there is no mis-measurement of background skill.[6] The model is only solved and analyzed in the two roommate case.

For two roommates i and j,

(1) $GPA_i = \delta + \alpha * ACA_i + \beta * ACA_j + \gamma * GPA_j + \varepsilon_i$

(2) $GPA_j = \delta + \alpha * ACA_j + \beta * ACA_i + \gamma * GPA_i + \varepsilon_j$

ε_i and $\varepsilon_j \sim N(0, \sigma_\varepsilon^2)$. By virtue of the random assignment of roommates, $E(\varepsilon_i , \varepsilon_j) = 0$.

Substituting (2) into (1) yields:

(3) $GPA_i = \delta + \alpha * ACA_i + \beta * ACA_j +$

$\gamma * (\delta + \alpha * ACA_j + \beta * ACA_i + \gamma * GPA_i + \varepsilon_j) + \varepsilon_i$

(4) $GPA_i = 1/(1-\gamma^2) * \left[\delta(1+\gamma) + (\alpha+\gamma\beta) * ACA_i + (\beta+\gamma\alpha) * ACA_j + \varepsilon_i + \gamma * \varepsilon_j\right]$

---

[5] I also include controls for answers to some housing questions and for gender as detailed in the next section. This lengthens the above equations somewhat, but the model works the same way as the simplified version shown.

Consider the OLS regression of GPA_i on ACA_i , ACA_j , and an intercept.

(5) $\quad E\,[\text{residual}] = E\left[1/(1-\gamma^2) * (\varepsilon_i + \gamma*\varepsilon_j)\right] = 0$

because ε_i , ε_j are independent and mean 0.

The OLS coefficients on ACA_i , ACA_j , and the intercept yield consistent estimates for $\delta(1+\gamma)$, $(\alpha+\gamma\beta)$, $(\beta+\gamma\alpha)$. Furthermore,

(6) $\quad \text{Var(residual)} = \text{Var}\left[1/(1-\gamma^2) * (\varepsilon_i + \gamma*\varepsilon_j)\right]$

$$= \frac{1}{(1-\gamma^2)^2}\,\text{Var}(\varepsilon_i + \gamma\varepsilon_i)$$

$$= \frac{1+\gamma^2}{(1-\gamma^2)^2}\,\sigma_\varepsilon^2$$

(7) $$\text{Cov}(\text{residual}_i\,,\ \text{residual}_j) = E\left[\left(\frac{\varepsilon_i}{1-\gamma^2} + \frac{\gamma\varepsilon_j}{1-\gamma^2}\right)\bullet\left(\frac{\varepsilon_j}{1-\gamma^2} + \frac{\gamma\varepsilon_i}{1-\gamma^2}\right)\right]$$

$$= \frac{2\gamma\sigma_\varepsilon^2}{1-\gamma^2}$$

[6] When this assumption is relaxed, the model is not identified.

The OLS coefficient estimates together with the variance and covariance of the residuals yield five equations allowing me to solve for the five unknowns which are α , β , γ , δ and σ_{ε}^{2} .[7]

In practice, I use ordinary least squares to estimate the reduced form given in (4) and I solve for the parameters. To obtain standard errors for the parameters, I use bootstrap samples to repeat the above procedure over and over, thereby generating an estimated distribution for each parameter. I use t-tests to check the significance of β and γ which are the effects of roommate observed background and roommate GPA respectively. This allows me to test the importance of roommate pre-treatment characteristics and roommate during-treatment outcomes.

A special case of the model occurs if I assume that the entire peer effect works through roommate outcomes and not background. (I.e. assume that roommate background does not enter in i's GPA directly. Equations (1) and (2) become

(8) $GPA_i = \delta + \alpha * ACA_i + \gamma * GPA_j + \varepsilon_i$

(9) $GPA_j = \delta + \alpha * ACA_j + \gamma * GPA_i + \varepsilon_j$

In this set-up, roommate background (ACA_j) is then an ideal instrument for roommate GPA_j because j's background is randomly assigned to i. Under this assumption, I can run two stage least squares to estimate the causal effect of j's GPA on i's GPA.

In addition to the above models, I also report results from a number of OLS and probit equations. For example, I show the simple OLS results from regressing i's outcomes on j's

---

[7] To ensure that the solution is unique I assume that $-1<\gamma<1$. This amounts to assuming that a 1.0 increase in j's GPA can not cause i's GPA to increase or decrease by more than 1.0. If γ were >1, any equilibrium would be unstable: a small increase in one roommate's GPA would cause both GPAs to go to infinity.

outcomes. These coefficients are subject to the endogeneity problem and can not be interpreted as causal. But I report them to show the amount of correlation in roommates' outcomes.

Data Set and Assignment Mechanism

The data come directly from Dartmouth's database of students. The data include a full history of housing/dorm assignments and term by term academic performance. Pre-treatment characteristics include SAT scores, HS class rank, public versus private high school, home state, and an academic index created by the admissions office. This last measure is constructed from test scores and high school grades adjusted for difficulty of high school program and competitiveness of high school. Outcomes include GPA, time to graduation, membership in fraternities, choice of major(s) and participation in athletics.

In addition, for the same students, I have more pre-treatment data from the Survey of Incoming Freshmen which is sponsored by the Higher Education Research Institute at UCLA. This is a survey of virtually all entering freshman across the US and provides me with a large set of pre-treatment characteristics, attitudes, and expectations. From the survey I use the following variables: parental income and education, student high school GPA, and whether or not the student reports drinking beer in the past year. I also have variables which capture the student's expectation about the likelihood of studying hard, graduating with honors, and joining a fraternity. The variables from the survey are available for at most 83% of my total sample. (Matching the data was only possible in cases in which survey respondents gave their social security number.) Some of the variables like "intention to join a frat" have a high rate of non-response.

Dartmouth freshmen are assigned to dorms and roommates randomly. Each freshman fills out (and mails in) a brief housing slip and the slips are then thoroughly shuffled by hand to create roommate groups which are then randomly assigned to dorms.

The assignment process is complicated by the fact that on the form each freshman answers yes or no to the following four statements: 1) I smoke (only 1% say yes to this); 2.) I like to listen to music while studying; 3.) I keep late hours; and 4) I am more neat than messy. Since rooms are separate by gender, this adds a fifth blocking variable for male versus female. The Office of Residential Life (ORL) groups the forms into 32 separate piles based on gender and the responses to the questions. Within each pile, the forms are shuffled by hand.

Then the piles are ordered randomly. There is a sheet for each different dorm and the sheet contains information on the available rooms. Each dorm is filled in the following manner: ORL takes dorm 1, room 1 and fills it with 1-3 students from pile 1 (depending on the room size). Then dorm 1, room 2 is filled from pile 2, and room 3 is filled from pile 3 and so on until dorm 1 is completely full. Subsequent dorms are filled in a similar manner until all of the freshman have been assigned to rooms and roommates. The effect of this process (as will be shown using the data) is to randomly assign students to dorms and to assign roommates who are random conditional on gender and the four housing questions.

ORL is "blocking" on the housing questions and this is the case that Rosenbaum and Rubin (1983) discuss in "Assignment to Treatment Group on the Basis of a Covariate." Conditional on the answers to the questions, the assignment is random. (In other words the assignment is random within a given block.) With the help of ORL, I retrieved all of the paper forms that the pre-freshmen had filled out. My research assistants then hand entered all of that data so that these key covariates would be available. Thus I am able to control for these pre-treatment covariates by measuring peer effects separately within each block.

In practice I do not actually show all of the analysis done block by block. In this specific case, it turns out to be possible to control for these covariates merely by including separate dummy variables for the answers to each question. This makes more efficient use of the available data. However, there are functional form assumptions inherent in this method of controlling for these important covariates. The analysis has also been done by blocking and is available upon request. The effects are all still present, though of course for some of the smaller blocks the t-stats are diminished.

The data used are for the classes of 1997 and 1998. I have data from several earlier classes, but these did not have random assignment of roommates. There was a policy change at ORL circa 1993 when the 97s were entering. Prior to the class of 97 there were several procedures which introduced a large amount of selection bias. Most importantly the housing forms contained a space for students to request a roommate and many students made these requests. Beyond that, ORL made some attempts to match together students who were thought to be both compatible and/or complementary. This was done mostly on any available information about home city, state, and country.

Within the classes of 97 and 98 there are still some people who make special requests for roommates, and I drop these people from the sample. For calculating the roommate variables, I use the original randomized freshman fall assignment. Only about 3% of people switch roommates during freshman year and ORL requires a strong reason to do so.

Summary Statistics

Table 1 contains summary statistics for the data.[8] Mean freshman year GPA is 3.20 and this tends to rise consistently throughout the sophomore, junior, and senior years reaching 3.40 in the senior year. Here I have calculated GPA independently for each year, rather than including the freshman grades in the sophomore GPA. Cohorts (classes) prior to the class of 1997 have similar numbers. In other words, GPA rises as students mature and/or take higher level classes. The GPA increase reflects this "time to graduation" effect as opposed to a general time trend in grade inflation.[9] Roommate 1 freshman year GPA has a mean of 3.21.

Roommate GPA is only defined where the freshman has one or more roommates, which is true for about 93% of the sample. The breakdown by room group size is as follows: 7.5% are in singles, 53% are in doubles, and the rest are in triples. In cases where there is more than 1 roommate, I average the data for the two roommates.[10]

Forty-nine percent of the sample is affiliated with a fraternity or sorority or co-ed Greek house. This is a binary variable which equals one if at some point during his or her Dartmouth career the student joined a fraternity. It need not have been during the traditional sophomore fall rush period and the student may have quit the organization at some point. Most fraternity members join sometime during their sophomore year and remain in the organization through graduation. The proportion joining a house is similar across men and women (not shown here). Currently I only examine this question as a binary outcome for membership. However, within fraternity members there is wide variation in the amount of time devoted to socializing, exercising, studying, and vacationing with fraternity/sorority brothers and sisters.

---

[8] I will go through these in detail because these variables (e.g. GPA, SAT scores) give the reader a good sense of the data and the outcomes being examined.

[9] I have five years' worth of data and do not find a grade inflation trend over that short period.

[10] In some of the analysis that follows, I show results just for rooms of two people. In particular, the estimation of the structural model requires this.

Only 3% of the sample graduates late. For these students this is defined as graduating any time after spring term senior year.[11] 11% of the students graduate as economics majors. The students are split roughly in thirds between majoring in the social sciences versus the natural sciences versus humanities. This is defined by their primary major. Double majors are allocated to the field that the student listed first on their major card. (Major is unknown for 4% of the sample.) Roughly 5% of the sample is black and 12% of all the students come from private high schools.

The mean math SAT is 691 and the mean verbal SAT is 631. The average class rank where known is 6. From the information on their pre-enrollment housing form, we see that 1% of the sample admits to smoking, 69% claim to be neat, 60% keep late hours, and 46% listen to music while studying. Certainly this self-reporting of behavior may not be 100% accurate. However, the potential for mis-reporting of behavior does not affect the ignorability of the assignment mechanism. Student i is equally likely to be assigned to any of the other students who gave the same answers. Note that when blocking on these covariates (the housing questions) the number of useful blocks is really at most 16 because almost no-one states that he/she is a smoker.

High school GPA is scaled as 1-8 where 8 is an A+ ; mean HS GPA is 7 which corresponds to an A. Father's and mother's education is scaled as 1-8. The mean of the variable is around a 6 for mothers which corresponds to college graduate.[12] The "drank beer" in the last year variable is coded as 1-3 corresponding to not at all, occasionally and frequently. 41% said not at all; 43% said occasionally and the rest said frequently.

[11] Almost all Dartmouth students entering as freshmen eventually graduate from Dartmouth, though some graduate late. These are often students who were sick, on suspension for academic or disciplinary reasons, or involved in extensive overseas programs or jobs.

Ignorability (randomness) of Assignment Mechanism

Table 2 shows that conditional on student i's responses to the housing questions, there is no relationship between i's pre-treatment characteristics and the pre-treatment characteristics of i's roommate. Regression 1 is an OLS of own math SAT score on roommate math SAT score and the housing questions. The t-statistic on roommate SAT score is -.61 indicating that there is no significant relationship among roommate math SATs, controlling for the housing question responses. Regressions 2,3,4 report similar results for verbal SAT score, HS academic score, and HS class rank. Note that for class rank we have fewer observations for which we have class ranks reported for both self and roommate.[13]

The responses to the housing questions are not particularly significant either. For example, in regression 1 which forms a linear predictor for math SAT, all of the t-stats are below 1.1. Being "neat rather than messy" raises the math SAT score by only 1.0 points, though it does appear to improve class rank.

The result of no relationship between roommate pre-treatment variables only holds in the classes for which ORL randomly assigned roommates. In regressions on some of the non-randomized data (not included) I find that roommate math SAT predicts own SAT with a t-statistic of 5.0.

---

[12] 7 is some graduate school and 8 is a graduate degree. In future drafts, it may be desirable to translate these codes into years of education. This would be roughly a linear transformation.

Peer Effects

Table 3A shows the results of regressing own outcomes on roommate outcomes and pre-treatment covariates including the housing questions. Since roommates are randomly assigned, the null hypothesis of no peer effects would predict no relationship between own outcomes and roommate outcomes.

In fact there is a significant relationship between own freshman year GPA and roommate freshman year GPA. Regression 1 shows this coefficient to be .11 with a t-stat of 4.3 controlling for own background and the housing questions. This implies that a 1.0 point increase in roommate GPA is associated with a .11 increase in own GPA. This effect is moderate in size and seems plausible given that we are dealing with students who have reached college age and have each already been heavily pre-screened for admission to Dartmouth.

Appendix 1 shows a similar regression in which I allow different slopes for the men and women. Here the slope for the women is .15. The slope for the men is the sum of the first two coefficients (the coefficients on roommate GPA and male*roommate GPA). The point estimate for the men is .08 which is 43% less than the slope for the women, though the difference is not statistically significant.

In Table 3A regression 1, <u>own</u> pre-treatment academic score has a coefficient of .014 and is highly significant (the t-stat is 14). This means that a 13 point increase in academic score (one standard deviation) raises freshman year GPA by about .18 or about 1/2 a standard deviation. Lower class rank (ie closer to number 1) is associated with higher freshman year GPA. But the coefficient is only -.001 which implies a small GPA effect for an improvement by 10 in class rank.

---

[13] The other way to run these regressions would be to include all of roommate pre-treatment covariates in each regression and report an F-statistic for the joint significance of all roommate pre-treatment variables. This yields

Table 3A, page 2 shows the coefficients on the housing questions. Smoking, keeping late hours, and listening to music are associated with lower GPA. The r-squared in regression 1 is .23, which indicates that my overall ability to explain differences in GPA using observables is somewhat limited.

Table 3A, regression 2 shows a probit of "member of fraternity/sorority" on freshman year roommate behavior and pre-treatment covariates. (Partials are reported rather than coefficients.) If my freshman year roommate joins a fraternity, I am 8% more likely to do so myself.[14] This is in spite of the fact that students do not even execute this decision during their freshmen year. Students are not allowed to join until sophomore year and only 16% of people keep any of the same roommates.

More remarkable is the frequency with which students join the same house as their randomly assigned roommate. Table 9 shows that fully 27% of roommate pairs who are both in fraternities join the same house. Under the null of no peer effect, this would be only 5% with a standard error of 1%.

Regression 3 in Table 3A shows that there is no significant relationship between own outcome and freshman year roommate outcome for "graduate late." This indicates that some key labor market outcomes may be completely unaffected by the types of peer effects for which I am testing. Regression 4 uses varsity athlete status as the outcome of interest and I run a probit of own participation in varsity athletics on roommate participation. The slope is basically zero.

Peer Effects in Choice of Major

---

similar results to those reported in Table 2.

[14] Unlike for GPA, the point estimates in column 2 are almost exactly the same if we run separate regressions for men and women or if we allow for different slopes and intercepts. Though some of the t-stats are less than 2. (See Appendix 1 for the different slopes regression).

A key manner in which roommates might affect long term labor market outcomes would be through student's choice of major. Choice of major or course of study has profound implications for eventual career choices and graduate school choices. However, the data show that randomly assigned roommates have no effect on choice of major.

Regressions 5 and 6 in Table 3A show probits of own major choice on roommate major choice. I find that roommate choice does not affect own choice significantly. For example in column 6, $\partial y/\partial x$ for own decision to major in one of the social sciences is .013 with a t-statistic of .23. This may be evidence against the information gathering version of social learning. If the peer effects worked mainly through information generated by the roommate's behavior, then one might expect major choice to be heavily influenced. As each roommate takes different classes, that should generate information which is useful to all members of the room.

Table 8 makes the same point about correlation in major choice utilizing a different statistical test. I compare the incidence of roommates with the same major against the incidence of "same major" that would be expected if major choices were randomly distributed across roommates. For example, since 36% of the students major within the humanities, under the null of no peer effects (i.e. under independence) one would expect 13% (.36*.36) of all roommate pairs to both be humanities majors. In fact, we do observe that 13% of the pairs are both humanities majors. The appropriate standard errors under the null hypothesis are also included in Table 8.

Table 3B goes looks at the same peer effects as Table 3A, but limits the sample to rooms where there are exactly two students. The results look similar to those in Table 3A. The coefficient of roommate freshman GPA on own freshman GPA is .14 which is similar to the coefficient in the larger sample. The peer effect on "frat" is about the same. A student is 9%

more likely to join a fraternity/sorority if her roommate does so. The peer effect on graduate late remains small and insignificant as does the effect on varsity athletic participation.

Basic Social Learning (Background) Versus During-Treatment Models

Table 4 shows estimates for the stuctural model. This is an attempt to remove the endogeneity problem in the estimates in Tables 3A, 3B. The estimates in Table 4 are intended to be estimates of the causal effects of roommate background and roommate outcomes.

In column 1, the coefficient on roommate GPA is .15 which is similar to the OLS estimate of .14 in Table 3B (which shows the two roommate case). The t-stat on roommate GPA is 1.6 for the model versus 4.4 under simple OLS. The coefficient on roommate HS academic index is small and insignificant under the structural model and under OLS. The implication is that while there is a significant peer effect, it does not work through roommate's background. Instead the peer effect works through the roommate's behavior and outcomes while at Dartmouth. This result is robust to using all my various measures of roommate pre-Dartmouth skill, eg SAT scores, HS GPA, parental education, and self-reported study habits.

Columns 2 and 3 of Table 4 run two stage least squares using a different set of assumptions. Here I assume that roommate background only affects own GPA indirectly. Hence the randomly assigned background characteristics can be used as instruments for roommate GPA. Using roommate academic index and SAT scores as an instrument (column 2), I find that the coefficient on roommate GPA falls to .04. When I include all possible roommate background characteristics as instruments (add family income, HS GPA, intent to study, intent to achieve honors, parent's education), I find that the IV coefficient rises to .28 and has a t-stat of 2.3.[15]

In column 4, I run the structural model to separate out roommate background from roommate outcome with regard to joining a frat. The most useful background variables for this outcome are family income and use of beer (pre-treatment). The results are similar in spirit to those for freshman GPA: My roommate joining a frat affects my outcome directly and does not work through my roommate's experience with beer. In the structural model, my roommate's decision to join a frat raises my likelihood of joining by 6%. In the IV formulation (assuming roommate beer does not enter directly), his joining a frat raises my likelihood of joining by 23%. This large increase in the coefficient (and insignificance) is probably due to the weakness of the instruments. The first stage r-squared is about .03. The results taken as a whole reject the basic social learning model in favor of the alternative models.

The Level of Aggregation

A further useful question is the level of aggregation at which the peer effects work. The data indicate that the fraternity membership effect works at the level of the entire dorm, whereas the GPA effect appears to work within a single room.

Table 10 shows the massive variation in fraternity participation by freshman dorm. This takes place despite the random assignment of dorms. For example, 15% of the 97s assigned to Cohen hall as freshmen eventually joined fraternities. This is statistically different than the class mean of 49%. However, 1 year later, the 98s assigned to that same hall as freshmen joined frats at a rate of 54%.

---

[15] The caveat to this last result is that the sample size falls to 377 due to non-response on the Survey of Incoming Freshmen.

These numbers are indicative of a dorm level peer effect. Social interaction among freshmen creates clumps of future fraternity members and non-members. This is similar to the social interactions model in Sacerdote, Glaeser, and Scheinkman (1996). The location of the clumps shifts from year to year as illustrated above with Cohen hall. This reinforces the idea that social interactions with dorm members are causing the agglomeration rather than location of the dorm or other fixed factors.

Table 5 addresses the level of aggregation question with several regressions. Column 1 shows a probit for frat membership. I include both average *roommate* frat membership and average frat membership on the student's whole *floor* as right hand side variables. This latter mean excludes own room. The effect ($\partial y/\partial x$) for *floor* average membership is .13 which is almost twice the partial of .07 on roommate membership.

In column 2, I increase the level of aggregation to look at the effect of average *dorm* membership (excluding own room) on a student's own frat membership. The effect of dorm behavior is .41 which is eight times larger than the effect of roommate behavior. These results tell a similar story to those of Table 10. The dorm level of frat membership is even more important than roommate behavior in determining whether or not a student joins a house.

Columns 3 and 4 of Table 5 tell a different story for GPA. For this outcome, neither average floor GPA nor average dorm GPA matters. But roommate GPA remains significant and has a coefficient of about .14 for the women and .08 for the men.

Can we identify the students who are most subject to peer effects?

The Survey of Incoming Freshman allows me to shed some light on this question. Prior to arrival, students were asked a battery of questions about the likelihood that they would engage in various activities including graduating with honors and joining a frat. Students responded that each outcome had either 1.) no chance, 2.) very little chance, 3.) some chance, or 4.) a very good chance.

In Table 6, I run separate regressions for the people who were unsure (responses 2,3) versus very sure (response 1,4). For example column 1 regresses own GPA on roommates' GPA for people who thought there was little chance or some chance that they would graduate with honors. The coefficient on roommate GPA is .11 which is much larger than the same coefficient for people who said there was no chance or a very good chance of graduating with honors. (Column 2 shows that the coefficient for these people is -.02.) This would indicate that people who were less certain about their outcome showed a much larger peer effect.

However, the results for fraternity membership (columns 3 & 4) are not as distinct. The effect for roommate frat on own frat is .19 if a student entered being unsure on this outcome. The peer effect only falls to .15 if a student entered with a strong conviction about this outcome. Both groups of students exhibit a large peer influence regardless of their initial convictions.

It is also true (results not shown here) that intention to join a frat is not a very good predictor of actual behavior. In contrast, intention to graduate with honors is a good predictor of GPA. The results indicate that there is some ability to use observables to determine who will be influenced by peers, but this clearly differs sharply depending on the outcome under consideration.

Who keeps their roommate?

Table 7 contains two probits examining who keeps their freshman year roommate into sophomore year. Men are 5% more likely than women to keep their roommates and students are 4% more likely to keep a roommate who is a member of a fraternity. In contrast, students are less likely to keep a roommate with a high academic index. The coefficient is -.002

which implies that a roommate with a 1 standard deviation higher (13.0) academic index is 2.6% less likely to be retained. Overall the coefficients and the pseudo R-squareds are small. The results may suggest that students with high socializing skills are valued as roommates slightly more than students with high academic skills.[16]

Peer Effects over Time

Figure 1 explores how the peer effect on freshman year GPA behaves over time. Here I plot the coefficients from regressing own GPA (in different time periods) on freshman year roommate GPA. The time periods do not cumulate; the "sophomore GPA" uses only grades from sophomore year as opposed to being the cumulative GPA. We see that the importance of the GPA peer effect from freshman year diminishes over time. By senior year, the effect diminishes to zero. Figure 1 shows both the "raw" coefficient and the coefficient controlling for own and roommate observables.

This attenuation of the peer effect could be explained in a variety of ways. One possible story is that as the students mix with each other more thoroughly over the four year period, the peer effect from the freshman year roommate becomes a less important component of total peer effects.

Conclusion

I find that roommate peer effects are important influences in freshman year GPA and in decisions to join social organizations. Roommate effects are not at all important in determining choice of major. The data reject a model of basic social learning from pre-treatment skills in

[16] Of course, its also possible that students with high academic skills are valued more highly and are more likely to be lured away into another group.

favor of a model that emphasizes during-Dartmouth behavior. The peer effect for fraternity membership is stronger at the dormitory level than at the individual room level.

Peer effects may be even more critical and long lasting earlier in student's lives (high school, junior high) and in a setting where there is more student heterogeneity. A fruitful area of future research could be to attempt to generate similar data in other settings.

Appendix

Comparison of Natural Experiment to OLS in Presence of Selection Problem

Following LaLonde (1986), Heckman (1998), and Dehejia and Wahba (1999), one could ask the following question: How do the treatment effects measured here under randomization compare to the effects that would be measured using standard econometric techniques in the presence of selection bias. For the treatment effects of job training programs, LaLonde finds that various econometric techniques are not successful in controlling for selection bias.

In the Dartmouth data, the selection biased and unbiased (randomized) coefficients are so close that I can not shed much additional light on the question . In Appendix 3 regression 1, I show my best estimate of the correct coefficient of roommate GPA on own GPA which is .11. Regression 2 shows a coefficient which is biased upward by selection. In regression 2, I do not control for answers to the housing form questions and I use data for the classes of 94-96. These classes contain extensive selection of roommates because they pre-date a housing office policy change as detailed above.

The coefficient in the selection biased regression is .14 which is 27% higher than the unbiased coefficient. However, in regression 3, I use OLS to attempt to control for the selection by including both own and roommate academic index. My OLS corrected coefficient is much closer to the "true" coefficient. Given that all three of the coefficients (true, selection biased, OLS corrected for bias) are close, little information can be gained about the ability of OLS to correct for selection bias in general.

In columns 4-6 I repeat the exercise for fraternity membership. The "true" partial is .08 for roommate frat on own frat. The selection biased effect is .14. Controlling for roommate background does nothing to reduce this gap. However, for the years 94-97, I do not have all relevant roommate background variables like pre-treatment "use of beer." It is possible that having more background variables would enable me to better correct for the selection bias in the coefficient.

References

Becker, G. (1997) "Social Economics," University of Chicago mimeograph.

Bernheim, D. (1994) "A Theory of Conformity," *Journal of Political Economy* 102 (5): 841-877.

Betts, J.R. and D. Morell (1999) "The Determinants of Undergraduate GPA: The relative Importance of Family Background, High School Resources, and Peer Group Effects," *Journal of Human Resources* 34: 268-93.

Borjas, G. (1995) "Ethnicity, Neighborhoods and Human Capital Externalities," *American Economic Review* 85: 365-390.

Brock, W. and S. Durlauf (1995) "Discrete Choice with Social Interactions I: Theory" NBER Working Paper # 2591.

Case, A. and L. Katz (1991) "The Company You Keep: The Effect of Family and Neighborhood on Disadvantaged Youth," NBER Working Paper # 3708.

Cochran, W.G. and D.B. Rubin (1973). "Controlling Bias in Observational Studies: Review." Sankhya- A, 35, pp.417-446.

Ellison, G. and D. Fudenberg (1993) "Rules of Thumb for Social Learning," *Journal of Political Economy* CI:612-643.

Ellison, G. and D. Fudenberg (1995) "Word-of-Mouth Communication and Social Learning" *Quarterly Journal of Economics* CX (1): 93-126.

Evans, W., Oates, W. and R. Schwab (1992) "Measuring Peer Group Effects: A Model of Teenage Behavior," *Journal of Political Economy* 100 (5) 966-991.

Gavaria, A. (1997) "Increasing Returns and the Evolution of Violent Crime: the Case of Columbia," UC San Diego Mimeograph.

Gavaria, A. and S. Raphael (1998) "School Based Peer Effects and Juvenile Behavior." UC San Diego Mimeograph.

Glaeser, E.L. and J. Scheinkman (1997) "Social Interactions and Long-Range Dependence"

Glaeser, E.L. and J. Scheinkman (1998) "Measuring Social Interactions"

Glaeser, E.L., Sacerdote, B. and J. Scheinkman (1996) "Crime and Social Interactions," *Quarterly Journal of Economics* CXI(2): 507-548.

Griliches, Z. (1958) "Research Costs and Social Returns: Hybrid Corn and Related Innovations," *Journal of Political Economy* 66: 419-431.

Hanushek E.A., J.F. Kain, and S.G. Rivkin (1998). "Do Higher Salaries Buy Better Teachers" (mimeo), University of Rochester.

Heckman J. and J. Hotz (1989). "Alternative Methods for Evaluating the Impact of Training Programs." Journal of the American Statistical Association.

LaLonde R. (1988). "Evaluating the Econometric Evaluations of Training Programs." American Economic Review.

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

D.B. Rubin (1974). "Estimating Causal Effects of Treatments in Randomized and Non-randomized Studies." Journal of Educational Psychology, 66, pp.688-701.

D.B. Rubin (1977). "Assignment to Treatment Group on the Basis of a Covariate." Journal of Educational Statistics, 2, pp.1-26.

Sah, R. (1991) "Social Osmosis and the Patterns of Crime," *Journal of Political Economy* XCIX, 1272-1295.

Young, P. (1993) "The Evolution of Conventions," *Econometrica* 61: 57-84.

Young, P. (1997) "Social Coordination and Social Change," Johns Hopkins mimeograph.

Table 1
Summary Statistics

Variable	*Obs*	*Mean*	*Std. Dev.*	*Min*	*Max*
freshman year GPA	1766	3.20	0.43	0.67	4.00
sophomore year GPA	1728	3.28	0.44	0.30	4.00
junior year GPA	1703	3.35	0.45	0.60	4.00
senior year GPA	1682	3.40	0.45	0.50	4.00
roommate freshman year GPA	1618	3.19	0.45	0.67	4.00
fraternity/sorority/co-ed house	1768	0.49	0.50	0.00	1.00
graduate late	1768	0.03	0.17	0.00	1.00
economics major	1768	0.11	0.31	0.00	1.00
social science major	1768	0.33	0.47	0.00	1.00
science major	1768	0.29	0.46	0.00	1.00
humanities major	1768	0.34	0.48	0.00	1.00
black	1768	0.05	0.22	0.00	1.00
SAT Math	1766	690.70	66.87	420.00	800.00
SAT Verbal	1766	631.26	71.66	360.00	800.00
academic score (incoming)	1736	203.87	13.14	151.00	232.00
high school class rank (incoming)	1768	5.54	10.48	0.00	75.00
high school class rank missing	1768	0.38	0.49	0.00	1.00
private high school	1768	0.12	0.33	0.00	1.00
smokes (housing form)	1768	0.01	0.11	0.00	1.00
more neat than messy (housing form)	1766	0.69	0.46	0.00	1.00
stays up late (housing form)	1767	0.60	0.49	0.00	1.00
listens to music (housing form)	1768	0.46	0.50	0.00	1.00
request substance free dorm (housing form)	1768	0.08	0.27	0.00	1.00
same roommate sophomore year	1768	0.16	0.36	0.00	1.00
father's education	1439	6.91	1.62	1.00	8.00
mother's education	1450	6.27	1.66	1.00	8.00
HS GPA	1464	7.39	0.84	2.00	8.00
Pre-Dart: drank beer in past year	1472	1.75	0.71	1.00	3.00
Pre-Dart: likelihood join frat	371	2.77	0.96	1.00	4.00
Pre-Dart: amount of time study	1245	5.34	1.56	1.00	8.00
Pre-Dart: likelihood play varsity	368	2.62	1.14	1.00	4.00
Pre-Dart: likelihood grad honors	405	3.06	0.64	1.00	4.00

Table 2
Own Pre-treatment Characteristics Regressed On Roommate Pre-treatment Characteristics

	(1) SAT Math (self)	*(2) SAT Verbal (self)*	*(3) HS Academic Index*	*(4) HS Class Rank*
roommate SAT math	-0.017			
	(-0.607)			
roommate SAT verbal		-0.001		
		(-0.046)		
roommate HS academic score			0.020	
			(0.736)	
roommate1 HS class rank				-0.018
				(-0.457)
smokes (housing form)	-15.341	6.202	-3.584	3.352
	(-1.091)	(0.411)	(-1.297)	(1.025)
more neat than messy (housing form)	1.049	-4.358	0.112	-1.097
	(0.294)	(-1.135)	(0.158)	(-1.297)
keep late hours (housing form)	-2.554	0.409	-0.751	-0.317
	(-0.738)	(0.110)	(-1.096)	(-0.393)
music while study (housing form)	-0.238	0.424	-0.469	0.670
	(-0.071)	(0.117)	(-0.706)	(0.839)
request substance free dorm (housing form)	6.334	14.548	3.155	-2.269
	(1.047)	(2.238)	(2.637)	(-1.733)
male	34.037	9.261	2.721	3.332
	(10.025)	(2.638)	(4.200)	(4.319)
constant	686.006	630.289	199.041	8.368
	(35.743)	(34.077)	(35.078)	(7.967)
R-squared	.06	.01	.02	.03
N	1610	1610	1591	999

T-statistics in parentheses
In cases with more than one roommate, roommate variables are averaged
Columns 1-4 are OLS.

Table 3A
Own Outcomes on Roommate Outcomes

	(1) Fresh year GPA	(2) Member frat/ soror	(3) Grad -uate late	(4) Varsity Athlete	(5) Econ Major	(6) Social Sciences Major
roommate freshman year GPA	0.110 (4.292)					
roommate member of fraternity/sorority/co-ed house		0.083 (2.860)				
roommate graduate late			0.010 (0.506)			
roommate varsity athlete				0.000 (-0.010)		
roommate is econ major					-0.078 (-1.810)	
roommate social science major						0.004 (0.110)
male	-0.080 (-4.069)	0.056 (2.130)	0.023 (2.483)	-0.002 (-0.120)	0.056 (3.880)	-0.016 (-0.680)
black	-0.033 (-0.640)	-0.281 (-4.260)	-0.013 (-0.575)	-0.043 (-0.910)	0.057 (1.240)	0.143 (2.380)
roommate black	0.016 (0.360)	0.078 (1.310)	-0.004 (-0.176)	-0.020 (-0.500)	-0.016 (-0.470)	-0.090 (-1.700)
HS academic score (self)	0.014 (14.733)	0.001 (0.400)	0.000 (-0.273)	0.000 (0.280)		
SAT math					0.001 (6.030)	0.000 (1.960)
SAT verbal					0.000 (-1.970)	-0.001 (-2.960)
HS class rank (self)	-0.001 (-0.820)	0.002 (1.390)	0.001 (1.656)	0.000 (-0.270)	-0.002 (-2.100)	0.000 (-0.050)
HS class rank missing (dummy)	-0.056 (-2.495)	0.135 (4.560)	0.018 (1.708)	0.023 (1.110)	-0.005 (-0.300)	0.029 (1.110)
private HS (self)	0.008 (0.254)	0.078 (1.790)	-0.013 (-0.854)	0.607 (15.130)	0.048 (1.900)	0.078 (2.040)
smokes (housing form)	-0.100 (-1.215)	0.059 (0.540)	0.005 (0.128)		-0.008 (-0.120)	0.006 (0.060)

more neat than messy (housing form)	0.039 (1.870)	-0.020 (-0.720)	0.004 (0.416)	0.023 (1.210)	0.014 (0.960)	0.023 (0.950)
keep late hours (housing form)	-0.059 (-2.869)	-0.015 (-0.540)	0.018 (1.926)	-0.035 (-1.860)	0.002 (0.140)	0.011 (0.470)
music while study (housing form)	-0.027 (-1.339)	-0.005 (-0.180)	-0.003 (-0.356)	0.008 (0.420)	-0.012 (-0.850)	-0.013 (-0.540)
request substance free dorm (housing form)	0.017 (0.468)	-0.164 (-3.430)	0.007 (0.412)	-0.074 (-2.260)	-0.047 (-1.940)	-0.071 (-1.700)
constant	0.053 (0.244)		0.023 (0.249)			
R-squared	.23	.04	.01	.25	.07	.01
N	1598	1602	1602	1580	1757	1757

T-statistics in parentheses
In cases with more than one roommate, roommate variables are averaged
Columns 2,4,5,6 are Probits. $\partial y/\partial x$ is shown. Columns 1,3 are OLS.

Table 3B
Own Outcomes on Roommate Outcomes
For Rooms W/ Two Students

	(1) Fresh year GPA	*(2) Member frat/ soror*	*(3) Grad -uate late*	*(4) Varsity Athlete*
roommate freshman year GPA	0.137 (4.433)			
roommate member of fraternity/sorority/co-ed house		0.088 (2.490)		
roommate graduate late			0.036 (1.070)	
roommate varsity athlete				0.020 (0.630)
male	-0.093 (-3.485)	0.032 (0.890)	0.008 (0.678)	-0.007 (-0.280)
black	-0.051 (-0.780)	-0.215 (-2.490)	-0.019 (-0.654)	-0.021 (-0.320)
roommate black	0.052 (0.875)	0.056 (0.710)	-0.003 (-0.097)	-0.016 (-0.290)
HS academic score (self)	0.014 (11.065)	0.001 (0.560)	0.001 (1.543)	0.001 (0.710)
HS class rank (self)	-0.002 (-1.347)	0.003 (1.600)	0.002 (2.533)	-0.001 (-0.570)
HS class rank missing (dummy)	-0.091 (-2.990)	0.148 (3.620)	0.032 (2.356)	0.009 (0.330)
private HS (self)	0.045 (1.053)	0.089 (1.550)	-0.024 (-1.239)	0.625 (11.560)
smokes (housing form)	-0.138 (-1.361)	0.092 (0.670)	0.035 (0.771)	
more neat than messy (housing form)	0.011 (0.406)	-0.052 (-1.400)	0.007 (0.549)	0.036 (1.450)
keep late hours (housing form)	-0.046 (-1.696)	-0.047 (-1.260)	0.012 (0.989)	-0.026 (-1.020)
music while study (housing form)	0.013 (0.503)	-0.040 (-1.110)	-0.006 (-0.489)	0.022 (0.880)

request substance free dorm (housing form)	0.051 (1.119)	-0.149 (-2.370)	-0.014 (-0.700)	-0.050 (-1.120)
constant	0.002 (0.006)		-0.179 (-1.494)	
R-squared	.26	.04	.02	.27
N	849	853	853	839

T-statistics in parentheses
In cases with more than one roommate, roommate variables are averaged
Columns 2,4 are Probits. $\partial y/\partial x$ is shown. Columns 1,3 are OLS.

Table 4
Peer Effects via Strutural Model, via I.V.

	(1) Fresh year GPA	*(2) Fresh year GPA 2sls*	*(3) Fresh year GPA 2sls*	*(4) Member frat/ soror*	*(5) Member frat/ soror 2sls*
roommate freshman year GPA	.148 (1.64)	0.040 (0.707)	0.282 (2.330)		
roommate member of fraternity/sorority/co-ed house				0.055 (2.607)	0.228 (1.008)
roommate HS academic index	-.003 (-1.3)				
HS academic score (self)	.014 (14.617)	0.014 (14.709)	0.017 (8.268)		0.000 (0.083)
roommate HS use of beer				0.013 (0.446)	
own HS use of beer				0.106 (3.796)	
male		-0.086 (-4.262)	0.030 (0.718)		0.059 (1.843)
black		-0.026 (-0.511)	0.009 (0.072)		-0.356 (-4.830)
roommate black		-0.005 (-0.103)	0.175 (1.497)		0.095 (1.210)
HS class rank (self)		-0.001 (-0.732)	0.002 (0.808)		0.003 (1.575)
HS class rank missing (dummy)		-0.058 (-2.540)	-0.055 (-1.173)		0.149 (4.606)
private HS (self)		0.010 (0.291)	0.004 (0.058)		0.043 (0.895)
smokes (housing form)		-0.108 (-1.299)	0.100 (0.550)		0.055 (0.427)
more neat than messy (housing form)		0.043 (2.014)	0.021 (0.449)		-0.006 (-0.196)
keep late hours (housing form)		-0.067 (-3.158)	-0.009 (-0.201)		-0.020 (-0.658)

music while study (housing form)		-0.028 (-1.398)	-0.017 (-0.392)		-0.048 (-1.635)
request substance free dorm (housing form)		0.018 (0.505)	-0.081 (-0.977)		-0.115 (-1.935)
constant		0.248 (0.921)	-1.224 (-2.078)		0.315 (0.999)
R-squared	.21	.22	.25	.03	.04
N	849	1587	377	628	1260

T-statistics in parentheses
In cases with more than one roommate, roommate variables are averaged
Columns 2,3,5 are two-stage least squares. In column 2, roommate HS academic index, SAT scores are used to instrument for roommate GPA. In 3, roommates' academic index, family income, HS gpa, intention to study, intention to graduate with honors are all used as instruments for roommate GPA. In 5, family income and HS use of beer are used to instrument for roommate decision to join frat.

Columns 1 and 4 are estimated via indirect least squares-- using the reduced form given in the text. T-stats are calculated by obtaining standard errors via bootstrapping.

Table 5
Level of Aggregation:
Room Versus Floor Versus Dorm Effects

	(1) Member frat/ soror	*(2) Member frat/ soror*	*(3) Fresh year GPA*	*(4) Fresh year GPA*
mean(frat) for **floor** exclud. own room	0.131 (1.630)			
mean(frat) for **dorm** exclud. own room		0.406 (2.910)		
roommate member of fraternity/sorority/co-ed house	0.074 (2.080)	0.060 (1.670)		
mean(fresh GPA) for **floor** exclud. own room			0.050 (0.722)	
mean(fresh GPA) for **floor** exclud. own room***male**			-0.052 (-0.491)	
mean(fresh GPA) for **dorm** exclud. own room				0.081 (0.420)
mean(fresh GPA) for dorm exclud. own room***male**				-0.136 (-0.612)
roommate freshman year GPA			0.146 (3.270)	0.144 (3.377)
male*roommate freshman year GPA			-0.061 (-0.995)	-0.057 (-0.840)
male	0.053 (2.220)	0.058 (2.410)	0.282 (0.721)	0.531 (0.696)
black	-0.279 (-4.210)	-0.277 (-3.190)	-0.034 (-0.633)	-0.035 (-0.608)
roommate black	0.071 (1.120)	0.069 (1.110)	0.020 (0.540)	0.020 (0.550)
HS academic score (self)	0.001 (0.460)	0.001 (0.450)	0.014 (14.987)	0.014 (13.566)
HS class rank (self)	0.002 (1.420)	0.002 (1.450)	-0.001 (-0.862)	-0.001 (-0.863)
HS class rank missing (dummy)	0.136 (3.970)	0.136 (4.100)	-0.055 (-2.630)	-0.056 (-2.872)
private HS (self)	0.077 (1.740)	0.079 (2.390)	0.009 (0.301)	0.008 (0.241)

smokes (housing form)	0.071	0.072	-0.097	-0.096
	(0.800)	(0.800)	(-1.059)	(-1.017)
more neat than messy (housing form)	-0.023	-0.022	0.040	0.040
	(-0.900)	(-1.050)	(1.969)	(1.970)
keep late hours (housing form)	-0.018	-0.018	-0.057	-0.058
	(-0.640)	(-0.720)	(-3.096)	(-3.222)
music while study (housing form)	-0.003	-0.002	-0.029	-0.028
	(-0.110)	(-0.060)	(-1.534)	(-1.583)
request substance free dorm (housing form)	-0.151	-0.141	0.017	0.017
	(-3.030)	(-3.030)	(0.489)	(0.492)
constant			-0.223	-0.305
			(-0.611)	(-0.407)
R-squared	.04	.05	.23	.23
N	1593	1602	1598	1598

T-statistics in parentheses. T-stats are corrected for clustering at the floor or dorm level.
Columns 1, 2 are probits with $\partial y/\partial x$.

Table 6
Interaction of Peer Effects w/ Own Background
Who is More Easily Influenced?

	(1) Fresh GPA | unsure about grad w/ honors pre-Dartmouth	*(2) Fresh year GPA | sure about grad w/ honors*	*(3) Member frat | unsure about joining pre-Dartmouth*	*(4) Member frat | sure about joining pre-Dartmouth*
roommate freshman year GPA	0.107	-0.017		
	(1.863)	(-0.155)		
roommate member of fraternity/sorority/co-ed house			0.189	0.149
			(2.330)	(1.400)
male	-0.108	-0.161	0.078	0.144
	(-2.392)	(-1.750)	(1.090)	(1.430)
black	-0.003	0.010	-0.339	0.201
	(-0.024)	(0.038)	(-1.350)	(0.820)
roommate black	-0.058	0.211	-0.177	0.030
	(-0.495)	(1.071)	(-0.820)	(0.140)
HS academic score (self)	0.014	0.022	-0.008	0.000
	(6.378)	(5.046)	(-2.160)	(0.060)
HS class rank (self)	0.001	0.002	-0.005	0.002
	(0.426)	(0.483)	(-1.020)	(0.530)
HS class rank missing (dummy)	-0.062	-0.175	0.245	0.209
	(-1.207)	(-1.659)	(2.940)	(1.930)
private HS (self)	0.080	0.088	0.140	-0.103
	(1.151)	(0.607)	(1.220)	(-0.600)
smokes (housing form)	-0.434		-0.399	
	(-1.941)		(-1.420)	
more neat than messy (housing form)	-0.027	0.271	0.031	-0.209
	(-0.561)	(2.577)	(0.410)	(-1.950)
keep late hours (housing form)	-0.029	-0.206	-0.112	0.042
	(-0.622)	(-2.171)	(-1.510)	(0.420)
music while study (housing form)	0.021	-0.001	0.000	-0.190
	(0.460)	(-0.011)	(0.000)	(-2.010)
request substance free dorm (housing form)	-0.078	0.034	-0.078	
	(-0.745)	(0.190)	(-0.550)	

constant	0.147	-1.274		
	(0.300)	(-1.395)		
R-squared	.23	.41	.11	.10
N	311	99	230	133

T-statistics in parentheses
In cases with more than one roommate, roomate variables are averaged
Columns 3,4 are Probits. $\partial y/\partial x$ is shown. Columns 1,2 are OLS.

Table 7
Probits for Keep Same Roommate

	(1) Keep Same Roommate	*(2) Keep Same Roommate*
roommate member of fraternity/sorority/co-ed	0.044	0.009
	(2.060)	(0.270)
roommate in one of 10 most popular frat/sorority (0-1)		0.045
		(1.320)
male	-0.046	-0.039
	(-2.440)	(-2.010)
roommate HS academic index	-0.002	-0.001
	(-1.710)	(-1.600)
roommate freshman year GPA	-0.006	-0.007
	(-0.190)	(-0.220)
roommate smokes	0.042	0.050
	(0.480)	(0.560)
roommate keep late hours	-0.021	-0.022
	(-1.080)	(-1.110)
roommate neat	-0.021	-0.022
	(-1.030)	(-1.060)
roommate music while study	-0.023	-0.022
	(-1.210)	(-1.160)
R-squared	.02	.02
N	1413	1413

T-statistics in parentheses.

Columns 1, 2 are probits with $\partial y/\partial x$ shown.

Table 8
Own and Roommate Major Choice Compared to Null Hypothesis of No Correlation In Major Choice

Bold shows fraction of sample in each cell
italics shows expected fraction if own choice and roommate choice are independent
(standard error under null is shown in parentheses)

	Roommate Division of Major			
	humanities	sciences	social sciences	total
Own division of major				
humanities	**0.13**	**0.11**	**0.12**	0.36
	0.13	*0.11*	*0.12*	
	(0.01)	*(0.01)*	*(0.01)*	
sciences	**0.10**	**0.10**	**0.10**	0.30
	0.11	*0.09*	*0.10*	
	(0.01)	*(0.01)*	*(0.01)*	
social sciences	**0.12**	**0.10**	**0.12**	0.34
	0.12	*0.11*	*0.11*	
	(0.01)	*(0.01)*	*(0.01)*	
total	0.25	0.21	0.24	1.00
N= 1,506				

Table 9
Frequency of Roommates Choosing Same Fraternity For Rooms of Two w/ Both Joining Frats

Fraction that Choose Same House	.27
(sd of average)	(.03)
N	230
Fraction Choosing Same House Under Null of Independent choice	.05
(sd of average)	(.01)

Fraction "same house" if independent is calculated using the proportion of students in each of 27 Houses. I assume that each of two roommates draws a house from the known (and uneven) distribution. I then calculate what fraction of roommates would end up in the same house under independence

Table 10
Agglomeration of Frat Membership Across Dorms

Dorm	mean(frat) 97s	N	mean(frat) 98s	N	Mean under null	Std Error under null	t-stat for 98s ~=.49
Butterfield	0.23	13	0.15	13	0.49	0.14	-2.42
Russell Sage	0.53	45	0.47	45	0.49	0.07	-0.31
Bissell	0.48	23	0.33	24	0.49	0.10	-1.54
Brown	0.65	17	0.50	18	0.49	0.12	0.08
Cohen	0.15	26	0.54	26	0.49	0.10	0.49
Little	0.57	30	0.42	24	0.49	0.10	-0.72
Fayerweather	0.51	35	0.30	33	0.49	0.09	-2.15
North Fayerweather	0.57	23	0.40	25	0.49	0.10	-0.90
South Fayerweather	0.68	25	0.60	20	0.49	0.11	0.98
Lord	0.59	27	0.56	32	0.49	0.09	0.82
Streeter	0.58	26	0.46	28	0.49	0.09	-0.27
Gile	0.43	40	0.43	44	0.49	0.08	-0.77
Massachusetts	0.47	36	0.61	33	0.49	0.09	1.33
North Massachusetts	0.71	28	0.58	31	0.49	0.09	1.01
South Massachusetts	0.55	29	0.45	33	0.49	0.09	-0.41
New Hampshire	0.41	46	0.23	43	0.49	0.08	-3.38
Topliff	0.48	52	0.54	52	0.49	0.07	0.70
Ripley	0.29	17	0.31	13	0.49	0.14	-1.31
Woodward	0.55	20	0.44	16	0.49	0.12	-0.42
Smith	0.59	17	0.50	16	0.49	0.12	0.08
French	0.68	37	0.41	37	0.49	0.08	-1.03
Hinman	0.59	37	0.44	41	0.49	0.08	-0.65
McLane	0.37	35	0.34	44	0.49	0.08	-1.98
Andres	0.51	35	0.67	39	0.49	0.08	2.21
Zimmerman	0.40	30	0.50	20	0.49	0.11	0.09
Morton	0.69	16	0.47	17	0.49	0.12	-0.16
Hitchcock	0.59	44	0.59	46	0.49	0.07	1.32
Wheeler	0.67	36	0.38	40	0.49	0.08	-1.45
Richardson	0.45	33	0.43	30	0.49	0.09	-0.62

Appendix 1
Own Outcomes on Roommate Outcomes
Separate Coefficients for Men and Women

	(1) Fresh year GPA	*(2) Member frat/ soror*	*(3) Grad -uate late*	*(4) Varsity Athlete*
roommate freshman year GPA	0.147 (3.706)			
male*roommate freshman year GPA	-0.062 (-1.221)			
roommate member of fraternity/sorority/co-ed house		0.074 (1.790)		
male*roommate frat		0.019 (0.320)		
roommate graduate late			-0.028 (-0.775)	
male*roommate graduate late			0.057 (1.298)	
roommate varsity athlete				-0.094 (-2.100)
male*roommate varsity athlete				0.154 (2.760)
male	0.118 (0.722)	0.046 (1.200)	0.020 (2.162)	-0.027 (-1.380)
black	-0.033 (-0.656)	-0.281 (-4.250)	-0.012 (-0.505)	-0.044 (-0.940)
roommate black	0.020 (0.446)	0.076 (1.280)	-0.003 (-0.164)	-0.025 (-0.630)
HS academic score (self)	0.014 (14.718)	0.001 (0.400)	0.000 (-0.236)	0.000 (0.400)
HS class rank (self)	-0.001 (-0.805)	0.002 (1.390)	0.001 (1.670)	0.000 (-0.180)
HS class rank missing (dummy)	-0.056 (-2.484)	0.135 (4.560)	0.018 (1.756)	0.025 (1.210)

private HS (self)	0.008	0.078	-0.013	0.606
	(0.229)	(1.780)	(-0.827)	(15.100)
smokes (housing form)	-0.097	0.059	0.007	
	(-1.174)	(0.540)	(0.180)	
more neat than messy (housing form)	0.040	-0.020	0.004	0.024
	(1.882)	(-0.710)	(0.382)	(1.290)
keep late hours (housing form)	-0.058	-0.014	0.019	-0.038
	(-2.807)	(-0.530)	(1.956)	(-2.020)
music while study (housing form)	-0.028	-0.005	-0.003	0.010
	(-1.393)	(-0.170)	(-0.376)	(0.530)
request substance free dorm (housing form)	0.017	-0.164	0.006	-0.072
	(0.487)	(-3.430)	(0.390)	(-2.220)
constant	-0.064		0.021	
	(-0.268)		(0.224)	
R-squared	.23	.04	.01	.26
N	1598	1602	1602	1580

T-statistics in parentheses
In cases with more than one roommate, roomate variables are averaged
Columns 2,4 are Probits. $\partial y/\partial x$ is shown. Columns 1,3 are OLS.

Appendix 2
Predicting Academic Score, GPA Using Pre-treatment Observables

	(1) HS Academic Index	*(2) HS Academic Index*	*(3) Fresh year GPA*	*(4) Fresh year GPA*
family income (14 categories)	0.286 (2.542)	0.118 (0.910)	0.003 (0.651)	0.004 (1.198)
Father's education (5 categories)		0.413 (1.475)	0.000 (0.041)	
Mother's education (5 categories)		0.446 (1.653)	0.006 (0.703)	
male			-0.072 (-3.142)	-0.069 (-2.917)
black			-0.460 (-8.524)	-0.108 (-1.707)
HS academic score (self)				0.012 (10.369)
HS GPA (self)				0.061 (4.293)
Intends to study hard in college				0.009 (1.238)
HS class rank				-0.002 (-1.813)
HS class rank missing				-0.060 (-2.164)
private HS (self)				0.026 (0.646)
smokes (housing form)			-0.196 (-1.852)	-0.116 (-1.098)
more neat than messy (housing form)			0.039 (1.535)	0.033 (1.285)
keep late hours (housing form)			-0.080 (-3.299)	-0.081 (-3.320)
music while study (housing form)			-0.044 (-1.827)	-0.034 (-1.415)

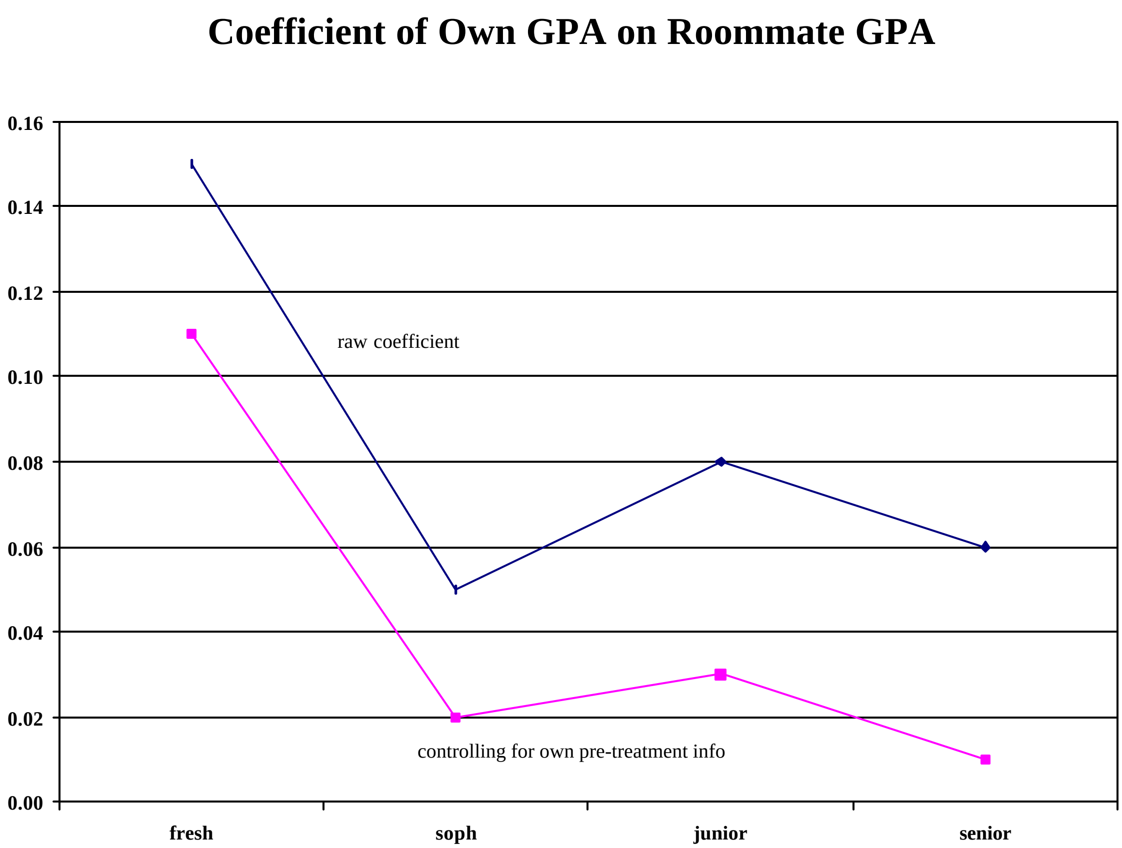
request substance free dorm (housing form)			0.064 (1.396)	0.015 (0.307)
constant	200.878 (162.141)	197.042 (113.324)	3.220 (50.121)	0.266 (1.008)
R-squared	.01	.01	.08	.22
N	1344	1324	1332	1137

T-statistics in parentheses
In cases with more than one roommate, roomate variables are averaged
Columns 1-4 are OLS.

App. 3: Ability of OLS to Control For Selection Problems

	(1) GPA sample w/ random room assign (1997-98)	(2) GPA sample w/ selection bias (1994-96)	(3) GPA sample w/ selection bias (1994-96)	(4) Frat sample w/ random room assign (1997-98)	(5) Frat sample w/ selection bias (1994-96)	(6) Frat sample w/ selection bias (1994-96)
roommate freshman year GPA	0.110 (4.292)	0.136 (6.483)	0.100 (4.775)			
roommate member of fraternity/sorority				0.083 (2.860)	0.140 (6.500)	0.140 (6.490)
male	-0.080 (-4.069)	-0.007 (-0.385)	-0.033 (-2.133)	0.056 (2.130)	0.090 (4.660)	0.088 (4.550)
black	-0.033 (-0.640)			-0.281 (-4.260)		
roommate black	0.016 (0.360)			0.078 (1.310)		
HS academic score (self)	0.014 (14.733)		0.015 (29.593)	0.001 (0.400)		0.000 (-0.420)
roommate HS academic score			0.000 (-0.746)			0.001 (1.950)
HS class rank (self)	-0.001 (-0.820)			0.002 (1.390)		
HS class rank missing (dummy)	-0.056 (-2.495)			0.135 (4.560)		
private HS (self)	0.008 (0.254)			0.078 (1.790)		
smokes (housing form)	-0.100 (-1.215)			0.059 (0.540)		
more neat than messy (housing form)	0.039 (1.870)			-0.020 (-0.720)		
keep late hours (housing form)	-0.059 (-2.869)			-0.015 (-0.540)		
music while study (housing form)	-0.027 (-1.339)			-0.005 (-0.180)		
request substance free dorm (housing form)	0.017 (0.468)			-0.164 (-3.430)		
constant	0.053 (0.244)	2.695 (40.385)	-0.014 (-0.100)			
R-squared	.22	.02	.26	.04	.02	.02
N	1598	2709	2709	1602	2715	2715

Figure 1



Coefficient of Own GPA on Roommate GPA
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
raw coefficient
controlling for own pre-treatment info
fresh
soph
junior
senior