

Does Your Cohort Matter? Measuring Peer Effects in College Achievement

Scott E. Carrell
UC Davis and NBER

Richard L. Fullerton
USAF Academy

James E. West
USAF Academy

April 14, 2009*

Abstract

We estimate peer effects in college achievement using a dataset in which individuals are exogenously assigned to peer groups of about 30 students with whom they are required to spend the majority of their time interacting. This feature enables us to estimate peer effects that are more comparable to changing the entire cohort of peers. Using this broad peer group, we measure academic peer effects of much larger magnitude than found in previous studies. The effects persist at a diminished rate into follow-on years and we find evidence of nonlinearities in the magnitude of the effects across student academic ability.

1 Introduction

Education policies such as integration, busing, and school choice often make substantial changes to the composition of a student's peer group. Likewise affirmative action in admissions provides

*The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the U.S. Air Force, Department of Defense, or the U.S. Government. Special thanks go to: William Carpenter, Rolland Stoneman, Kathleen O'Donnell, Jau Tsau, Richard Schreiner and Kate Carson for their assistance in obtaining the data and background information for this project. Thanks also go to Josh Angrist, Sandra Black, Pascaline Dupas, Caroline Hoxby, Changhui Kang, Lawrence Katz, Marianne Page, Bruce Sacerdote, Andrew Samwick, Doug Staiger, Omari Swinton, and all seminar participants at the NBER Summer Institute, WEA Meetings, Baylor University, Dartmouth College, Florida State University, Georgia State University, Ohio State University, Rand Corporation, Rice University, Sam Houston State University, Syracuse University, Texas A&M, U C Davis, University of Florida, University of Houston, and University of Texas-Austin for their helpful comments and suggestions.

Carrell: (530)302-1038 (v) secarrell@ucdavis.edu

Fullerton: (719)333-3080 (v) richard.fullerton@usafa.edu

West: (719)333-8735 (v) jim.west@usafa.edu

opportunities for greater access to more selective colleges and universities for traditionally under-represented groups. Such policies are largely predicated upon the assumption of significant positive peer effects in educational outcomes. Despite this, relatively little is known about how the quality of one’s *entire* cohort of peers actually affects individual outcomes. That is, how would student outcomes differ by having a cohort of higher quality peers, all else equal?

To date, the most convincing postsecondary peer effects studies have exploited situations where students have been randomly assigned to roommates and/or dorms.¹ Results from these studies have found only mixed evidence regarding the existence of positive peer effects in academic performance.² A major drawback of these studies, however, is that roommates are generally only a small subset of an individual’s actual peer group.³ Thus, works in the previous literature have likely underestimated the total magnitude of peer effects due to measurement error in the peer group. Additionally, “roommate” studies provide relatively little information regarding how an individual’s outcomes would differ if his/her entire cohort of peers were to change, as would be the case if the individual were to attend a different college.

In this study, we exploit a unique dataset in which individuals are exogenously assigned to peer groups of about 30 students, with whom they are required to spend the majority of their time interacting. Conditional on a few demographic characteristics, the students in our study are randomly assigned to a peer group in which they live in adjacent dorm rooms, dine together, compete in intramural sports together, and study together. They have limited ability to interact with other students outside of their assigned peer group during their freshman year of study. This feature enables us to estimate peer effects that are more comparable to changing the entire cohort of peers. Students are also randomly assigned to roommates within the peer group, which allows us to make comparisons with the previous “roommate” peer effects literature.

Our results are significant for several reasons. First, we find academic peer effects that are

¹The major exception is Lyle (2007) who examines peer effects using data from the US Military Academy (USMA).

²To date there has been little evidence of large positive peer effects in academic performance. Sacerdote (2001) finds evidence of small contemporaneous peer effects for Dartmouth roommates. Zimmerman (2003) finds small roommate contextual effects for individuals in the middle 70-percent of the distribution at Williams. Foster (2006) and Lyle (2007) find no evidence contextual peer effects at Maryland and West Point. Stinebrickner and Stinebrickner (2006) finds no evidence of peer effects for males and small contextual effects for females at Berea College. Also see: Hoxby and Weingarth (2006), Siegfried and Gleason (2006), Li and Li (2009) and Kremer and Levy (2008). Non-academic studies such as Falk and Ichino (2006) and Mas and Moretti. (2009) have found evidence of social spillovers in task oriented settings.

³Evidence suggests that college students quickly establish networks of friends and study partners that extend beyond the roommate or dorm level (Stinebrickner and Stinebrickner, 2006).

much larger in magnitude than the previous literature when we estimate the effects using the broad peer group to which individuals are assigned. For freshman students, a 100-point increase in the peer group average SAT verbal score increases individual GPA by roughly 0.4 grade points on a 4.0 scale. Second, using course-level data, we find that peer effects are largest in math and science courses and virtually non-existent in physical education and foreign language courses. Because physical education and foreign language courses have the least opportunities for interaction among students, these findings suggest that peer effects may be working through study partnerships versus a social norm of effort.⁴ Third, we find the freshman peer effects persist at a diminished rate into the sophomore, junior, and senior years. These results show that social network peer effects during the freshman year significantly affect achievement in follow-on years. Finally, we find evidence of nonlinearities in the magnitude of the peer effects across student incoming academic ability. These results imply there may be social gains to sorting students across peer groups.

Our results also help explain why many previous studies have found little evidence of academic peer effects in higher education: the bulk of those studies focus on roommates and dorm floors. We find that roommates and dorm floors capture only a limited proportion of an individual's peer group. Like Sacerdote (2001) and Zimmerman (2003), we find only moderate evidence of peer influence at the roommate level. Defining the peer group as the set of students who live in close geographic proximity, as in Foster (2006), we do not find measurable peer effects.

The remainder of the paper unfolds as follows. Section 2 reviews the challenges in measuring peer effects and describes the strategy we use. Section 3 explains the peer group structure in our study. Section 4 describes the data. Section 5 presents our statistical methods and results. Section 6 concludes.

2 Measuring Peer Effects

Manski (1993) distinguishes three types of peer influence: endogenous effects, exogenous effects, and correlated effects. Measuring the importance of each of these effects is difficult for two main reasons. First, it is often difficult to separate the effect that the group has on the individual from the effect the individual has on the group because outcomes are jointly determined (Vigdor and Nechyba, 2007). This problem is often referred to as the endogeneity problem (Moffitt, 2001; Sacerdote,

⁴Not all students are required to take a foreign language and students are spread across foreign language courses. Therefore, the likelihood of finding a suitable study partner within a given peer group is much smaller than for other freshman courses, which have near universal common enrollment.

2001) or the reflection problem (Manski, 1993). The second issue in measuring peer influence occurs because individuals tend to self-select into peer groups. In the presence of self-selection, it is difficult to distinguish the peer effects from the selection effects (Sacerdote, 2001).

The endogeneity problem is typically handled by finding suitable instruments for peer behavior that are exogenous to the stochastic error component of the dependent variable. A more recent strategy in the education peer effects literature is to use previous peer achievement as an instrument for current achievement (Betts and Zau, 2004; Burke and Sass, 2008; Hanushek et al., 2003; Vigdor and Nechyba, 2007). While this strategy is presumably used as a consequence of data constraints, lagged peer achievement may not be exogenous to contemporaneous achievement.⁵

The selection problem has been handled in two main ways. A first strategy (widely used in the primary education peer effects literature) is to exploit variation across classrooms or cohorts within a school.⁶ This has typically been accomplished using large administrative panel datasets while employing individual, school, grade, and/or year fixed effects. The second strategy, used by a growing literature measuring peer effects in higher education, is to exploit situations where individuals are randomly assigned to peer groups.⁷

In this paper, we use the random assignment of students at the United States Air Force Academy (USAFA) to broad social-network peer groups, called squadrons, as the main source of identification of peer effects. Our analysis builds on the previous literature in several ways. First, the randomization process at the USAFA allows us to measure peer effects at multiple peer group levels: roommate pairs, classmates within the same squadron, and upper classmen within the squadron. Second, we can use the vast amount of exogenous pre-treatment data to correct for endogeneity and measure peer effects using various measures of peer academic and non-academic ability. Third, students at USAFA take a set of approximately 30 mandatory core courses. We use these feature to estimate the peer effects across different types of courses free from selection bias into or out of course or section.⁸ Finally, reassignment to a new squadron peer group at the beginning of the sophomore year enables us to test for persistence in peer effects over time. In general, we find strong, robust peer effects of larger magnitude than those found in previous studies such as Sacerdote (2001),

⁵This is because many of the peers in an individual’s current peer group were also likely to be peers in the previous period(s). Hence, previous peer achievement is not exogenous to individual current achievement due to the cumulative nature of the education production function.

⁶See: Carrell, Malmstrom, and West (2008), Hoxby and Weingarh (2006), Vigdor and Nechyba (2007), Betts and Zau (2004), Burke and Sass (2008), Hanushek et al. (2003), among others.

⁷See: Boozer and Cacciola (2001), Foster (2006), Sacerdote (2001), Zimmerman (2003), Lyle (2007), among others.

⁸Which section of core courses students are enrolled in is determined by the USAFA Registrar without any input from students. “Convenience changes” are not allowed.

Zimmerman (2003), Foster (2006), and Stinebrickner and Stinebrickner (2006).

Lyle (2007) studies a similar peer-group setting at the US Military Academy (USMA). Our study is distinct from Lyle (2007) in both scope and the process in which students are assigned to peer groups. Lyle (2007) primarily focuses on biases due to common shocks in estimating contemporaneous peer effects models.⁹ Peer groups at the USMA are constructed to have an even distribution of student academic ability. This leveling of aptitude across peer groups reduces the exogenous variation of pre-treatment characteristics across groups, making it difficult to estimate reduced-form peer effects models.¹⁰

We recognize that questions could be raised about the generalizeability of our findings given USAFA students are a subset of traditional college students. While 17 percent of Air Force Academy students have parents who served in the military, the student body as a whole is drawn from the same pool as other selective academic institutions throughout the United States. In economic experiments to investigate behavior in real and hypothetical referenda, Burton et al. (2007) finds the behavior of students at USAFA and Queen’s University, Belfast to be statistically indistinguishable.

Because students at USAFA are taught to foster teamwork, our peer effects estimates could be larger than those expected at other institutions. However, institutional social constraints at USAFA (i.e., mandatory study periods, inability to attend fraternity parties, and big penalties for underage drinking) may result in smaller counterproductive peer influences. If true, properly measured peer groups in other institutional settings could exhibit larger peer effects than we find at USAFA. Further information regarding peer group formation at other institutions would be required to empirically test which effect dominates.

3 Peer Group Assignments at the Air Force Academy: A Natural Experiment

The Air Force Academy is a fully accredited undergraduate institution of higher education with an approximate enrollment of 4,500 students. There are 32 majors offered in the humanities, social sciences, basic sciences, and engineering. The average SAT for the 2005 entering class was 1,309 with an average high school GPA of 3.60 (Princeton Review, 2007). Applicants are selected for admission on the basis of academic, athletic, and leadership potential. Students are grouped into

⁹Contemporaneous models regress individual outcomes, such as GPA, on peer group outcomes.

¹⁰USAFA, by contrast, does not attempt to level academic ability across groups. As evidence, the variance of peer group-level SAT scores at USAFA is 49-percent greater in our data compared to Lyle (2007)’s USMA data.

one of 36 peer groups, called squadrons, with each group comprised of approximately 120 students (freshman through seniors). For their first seven months in the academy (from September through the end of March), freshman students are not allowed to enter the premises of another squadron. Hence, interaction with students from other squadrons is extremely limited for the freshman.¹¹

A significant amount of academic and social interaction takes place among students within each squadron. This forms a solid foundation to measure the “total peer effect” (Sacerdote, 2001) or total social influence for each individual. As freshman students are probationary members of a squadron, we would expect the primary peer group of freshman students to be that of other freshman students within the same squadron. However it is plausible that more senior members of a squadron could provide academic assistance as well as being mentors and leaders to the freshmen.

Measuring peer effects among USAFA students is simplified by the way the Academy places students in squadrons. The USAFA admissions office implements a stratified random assignment process where females are first randomly assigned to squadrons. Next, male ethnic and racial minorities are randomly assigned, followed by male non-minority recruited athletes. Students who attended a military preparatory school are then randomly assigned. Finally, all remaining students are randomly assigned to squadrons.¹² This stratified process is accomplished to ensure demographic diversity across peer groups. Roommate assignments are likewise made without student input. This structure creates a natural experiment for estimating peer influence. The overwhelming majority of entering students do not know anybody currently enrolled at USAFA. The appointment process, by which each member of the U.S. Congress and Senate nominate candidates from their congressional district or state, insures geographic diversity.

In attempting to develop an ability to work with peers of all abilities and backgrounds, USAFA does not ask any questions of incoming students as to their likes, dislikes, or roommate preferences. One might argue that the effect the institution is trying to achieve in bypassing student preferences (and for us, self-selection bias) is a behavioral model similar to the Rainbow model outlined in Hoxby and Weingarth (2006) where students benefit from interacting with all types of peers.

¹¹Freshman students can interact with students from other squadrons in academic classes, the library, gym, religious services, and what would be considered the student union. Freshman students who are on intercollegiate athletic teams or participate in club sports interact with students from other squadrons during practice times and on team trips.

¹²Students with the same last name, including siblings, are not placed in the same squadron.

4 Data

4.1 The Dataset

Our dataset includes all students in the USAFA graduating classes of 2005 through 2007. The students at USAFA are relatively high achieving, with the average SAT math scores in approximately the 88th percentile compared to nationwide SAT test takers.¹³ The 25th and 75th percentile distributions of SAT scores at USAFA are very similar to some of the top public universities in the United States such as UCLA, University of Michigan, University of Virginia, and UNC-Chapel Hill.¹⁴ Eighteen percent of the sample is female, 5-percent is black, 6-percent is Hispanic and 5-percent is Asian. Twenty-seven percent are recruited athletes and 2-percent attended a military preparatory school. Table 1 provides a complete list of summary statistics.

Pre-Academy (pre-treatment) data include whether students were recruited as athletes, whether they attended a military preparatory school, and measures of their academic, athletic and leadership aptitude.¹⁵ Our outcome performance data contain each individual's freshman through senior academic performance as measured by their grade in each course computed on a zero to 4.0 scale. Grades are determined on an A, A-, B+, B ··· C-, D, F scale where an A is worth 4 grade points, an A- is 3.7 grade points, a B+ is 3.3 grade points, etc.¹⁶

Grades are a consistent measure of academic performance across all students in our sample, since students at USAFA spend their entire freshman year taking required core courses and do not select their own coursework. The USAFA Registrar generates the fall semester academic schedules for the freshmen without any input from the affected students (the one exception is the choice of the foreign language requirement).¹⁷ Students have no ability to choose their professors. Core courses are taught in small sections of around 20 students, with students from all squadrons mixed across classrooms. Faculty teaching the same course use an identical syllabus and give the same exams during a common testing period. Grades for each course by semester are determined on

¹³See http://professionals.collegeboard.com/profdownload/sat_percentile_ranks_2008.pdf for SAT distributions.

¹⁴For our sample the 25th and 75th percentile SAT scores were 620 and 700 for math and 590 and 670 for verbal. For SAT distributions for top public universities see: http://collegeapps.about.com/od/sat/a/SAT_Public_Univ.htm.

¹⁵In addition to SAT scores, our data set includes for each student an academic composite (high school GPA, class rank, quality of school, and size of school), a leadership composite (high school and community activities), and a fitness test score.

¹⁶There are no explicit grading curves at USAFA. Therefore, a benefit to one student or group is not, mechanically at the expense of another student or group.

¹⁷Carrell and West (2008) show that course section placement is effectively random at USAFA conditional on an even distribution of females and athletes across sections within a course.

the same grading scale for all students in the course, regardless of instructor.¹⁸ These institutional characteristics assure there is no self-selection of students into courses or towards certain professors.

Absence of self-selection into courses or professors allows us to rule out potential mechanisms driving our results. First, peers influencing the choice of courses, professors, or academic major cannot be responsible for our results. Second, because students from all squadrons are randomly mixed across classrooms, our results are not driven by classroom interactions or common shocks.¹⁹

4.2 Are Peer Group and Freshman Roommate Assignments Truly Random?

Across the three years of our sample, there are 108 separate squadron peer groups with an average of 32.7 freshman students. To properly identify peer effects, there must be sufficient variation in the pre-treatment characteristics across groups. Under pure random assignment, the standard deviation of each peer group attribute should be equal to the population standard deviation divided by the square root of 32.7. This is largely the case in our sample. For example, the standard deviation of peer group SAT verbal score is 11.4 where population standard deviation is 11.7 ($67.0/\sqrt{32.7}$).

We were not able to find any official USAFA records for freshman roommate assignment. However, using a log of issuing and returning dorm room keys, we were able to successfully match approximately two-thirds of freshman students as roommates. We considered individuals as roommates if students were issued a key to the same room for a minimum of two overlapping months.

Of prime importance to our study is that students are not placed into squadrons or with (freshman) roommates based on pre-treatment performance. For each graduating class, we test for randomness in the squadron and roommate assignments in Table 2, which shows how individual pre-treatment characteristics are correlated with roommate and squadron pre-treatment characteristics (*academic composite*, *SAT math*, *SAT verbal*, *fitness score*, and *leadership composite*). Section A shows results with no controls, while Section B includes randomization control indicators for female, black, Hispanic, athlete and preparatory school.²⁰ Results indicate little evidence of systematic correlation across peer groups in the data. Of the 90 estimated coefficients, only six are statistically

¹⁸In some core courses, 5 to 10-percent of the overall course grade is earned by professor/section specific quizzes and/or class participation.

¹⁹Examples of common shocks could be a shared professor, questions raised in the classroom, or a barking dog outside the classroom. To completely rule out any of these possibilities, we include a course by section fixed effect in all of our empirical models.

²⁰As suggested by Guryan, Kroft, and Notowidigdo (2007) in Section B we also control for the average characteristics of possible peers in the roommate specifications to correct for a mechanical negative bias.

significant at the 0.10-level.²¹

5 Methods and Results

We first analyze peer effects using a reduced form linear-in-means model where we regress individual outcomes on roommate and peer pre-treatment characteristics.

Specifically, we estimate the following equation for academic performance:

$$G_{isct} = \phi_0 + \phi_1 X_{ist}^r + \phi_2 \frac{\sum_{k \neq i} X_{kst}}{n_{st} - 1} + \beta X_{ist} + \gamma_{ct} + \epsilon_{isct} \quad (1)$$

where G_{isct} is the freshman fall semester course grade for individual i in squadron s , in course-section c and in semester-year t . X_{ist}^r are the pre-treatment characteristics of individual i 's roommate²² and $\frac{\sum_{k \neq i} X_{kst}}{n_{st} - 1}$ are the average pre-treatment characteristics of all other peers in squadron s except individual i . X_{ist} is a vector of individual i 's specific (pre-treatment) characteristics, including SAT math, SAT verbal, academic composite, fitness score, leadership composite, race/ethnicity, gender, recruited athlete status and whether they attended a military preparatory school. γ_{ct} are course-section by semester-year fixed effects, which control for unobserved mean differences in academic achievement across courses, sections within courses, and time. ϵ_{isct} is the error term. Given the potential for error correlation across individuals within a given squadron and class, we correct all standard errors to reflect clustering at the squadron by class level.

5.1 Main Results

We estimate various specifications of (1) using ordinary least squares for freshman academic performance, with results shown in Table 3.²³ For Specification 1, we estimate the peer influence at the roommate level using the full array of roommate-level academic, athletic, and leadership

²¹Nine coefficients should be found significant at the 0.10-level with random sampling.

²²Twenty-nine percent of our sample were assigned to rooms with three students. For these students, the average pre-treatment characteristics of their two roommates are used as explanatory variables. We imputed missing roommate pre-treatment data using the mean of each variable in squadron s and all regressions include a dummy variable for missing roommates. As an alternative, we imputed missing roommate data using the mean of students in squadron s for whom we could not identify a roommate using the key log. Using this alternative specification, estimated coefficients varied by less than 0.001. Estimated coefficients are qualitatively similar when dropping students with missing roommates from the sample.

²³SAT scores, academic composite, leadership composite and fitness scores are divided by 100 prior to estimation.

pre-treatment measures,²⁴ which we find jointly insignificant. Own SAT verbal (0.089), SAT math (0.230), academic composite (0.110) and fitness score (0.052) are all positive and highly significant. The own leadership composite is positive and statistically insignificant.

For Specification 2, we estimate the model using the average pre-treatment characteristics of individual i 's peers (other freshmen) in squadron s . Of the five peer variables estimated, two coefficients are statistically significant, peer SAT verbal (0.338) and peer fitness score (0.153). Compared to previous studies the magnitude of peer SAT verbal is quite large. Similar to Zimmerman (2003), the reduced form academic peer effect appears to be driven through SAT verbal scores versus other academic pre-treatment measures.

Next, we estimate Specification 3 using the average pre-treatment characteristics of the three upper classes in the squadron to measure the effects from the upperclassmen within the squadron. Of the 15 upper class variables estimated, the sophomore class SAT verbal (0.227) and junior class fitness score (0.150) and leadership composite (0.049) are individually significant. All fifteen variables are jointly significant at the 0.01-level. This result implies that the characteristics of upperclassmen, as a whole, play an important role in freshman academic performance.

In Specification 4 we estimate the model using all peer and upper class pre-treatment characteristics. The model shows that both the peer and upper class pre-treatment characteristics are jointly significant at the 0.01-level. In Specification 5 we estimate the model using all pre-treatment characteristics for roommates, peers, and upperclassmen. In total, we estimate 25 different effects with 5 each for roommate(s), peers, sophomores, juniors, and seniors within the squadron. Overall, there are five positive and statistically significant coefficients: 1) roommate leadership composite (0.011), 2) peer SAT verbal (0.416), 3) peer fitness score (0.166), 4) sophomore class SAT verbal (0.259), and 5) junior class leadership composite (0.068). All 25 estimated coefficients are jointly significant at the 0.01-level (F -statistic p -value = 0.0003).

Finally, in Specification 6, we estimate a restricted model excluding peer SAT math, academic composite, and leadership composite due to their likely collinearity. The coefficients for the peer SAT verbal (0.382) and peer fitness (0.145) remain highly significant with similar magnitudes.

The previous results provide strong evidence of positive social spillovers in academic performance. As in Zimmerman (2003), we find the peer effects are linked more closely with SAT verbal scores versus other academic pre-treatment measures. These results also show that other non-academic measures, such as the athletic and leadership measures, are linked with positive peer

²⁴For student who only have a reported ACT score, we converted the ACT scores to SAT scores using conversions from the College Board (Dorans, 1999).

influence. The small roommate effects are consistent with previous studies, while the large positive peer effects at the squadron-level highlight the importance of properly identifying the relevant peer group when estimating peer effects.²⁵ The model estimates that a one-standard deviation increase in peer SAT verbal results in an increased own GPA of 0.05 grade points (one-twelfth of a standard deviation). In terms of standard deviations, this effect is nearly 2.5 times greater in magnitude than that found by Zimmerman (2003) for roommates at Williams College.²⁶

One could speculate that these large peer effects are purely driven by the institutional nature of USAFA (i.e., the military setting fosters more teamwork). However, the small roommate effects are not consistent with that hypothesis. That is, if military organizations were more affected by peer influence, we would also expect to see larger peer effects at the roommate-level compared to previous studies. The absence of large effects at the roommate-level indicates the institutional setting at USAFA is not solely responsible for the results.

5.2 Differences Across Types of Courses

Students at USAFA are required to take a core set of approximately 30 courses in mathematics, basic sciences, social sciences, humanities, and engineering throughout their four years of study. We use these common set of courses to examine the peer effects across course types during the entire freshman year free from selection bias into our out of courses.

Table 4 presents results for this analysis. For comparison, Specification 1 shows the full sample results from Table 3, Specification 5. Specification 2 shows results for math and science courses. The estimated coefficient for the peer SAT verbal variable (0.509) is large and highly significant, 22-percent larger than in the full sample. The model estimates that a 1-standard deviation increase in peer SAT verbal increases math and science performance by 0.06 grade points. Specification 3 shows results for humanities and social science courses. Again the peer SAT verbal variable is positive and statistically significant (0.405) though smaller than in the math and science courses.

Specifications 4 and 5 present results for foreign language and physical education courses. In both specifications there is little evidence of peer influence. The foreign language results are not surprising as students who take a foreign language are spread across various languages, limiting

²⁵In results not shown, we also tested whether the effects were larger for peers within the squadron with a closer demographic distance (i.e., females, minorities, and athletes). We found no evidence to support this hypothesis.

²⁶Zimmerman (2003) found that a 100-point increase in roommate SAT verbal increased own GPA by 0.03 grade points (Table 3) and a 1-standard deviation increase in roommate SAT verbal results in a 0.022 increase in own GPA.

the opportunity for peer interaction within a squadron. Results for the physical education courses show a small positive effect for peer fitness scores (0.116).

Finally, Specification 6 shows results for the military studies courses. The peer SAT verbal variable is positive and statistically insignificant (0.146) and the peer fitness (0.154) and the peer leadership (0.122) variables are statistically significant.

The preceding findings by subject suggest the peer effects may be working through study partnerships versus a social norm of effort. We suggest this because physical education and foreign language courses have the least opportunities for interaction among students. But we recognize that this hypothesis is not testable within our models as specified. It is also interesting that peer SAT verbal scores, versus peer SAT math scores, affect student math and science grades. Again, we can only speculate, but after informal discussions with faculty and students at the USAFA, one hypothesis is that students with high SAT verbal scores may exhibit better communication skills and a willingness to participate in study partnerships.

5.3 Persistence of the Effects

With evidence of positive peer effects in freshman academic performance, we next examine the persistence of these effects in subsequent years. It is possible to isolate freshman peer effects from follow-on peer effects because all students are (conditionally) randomly assigned to a new squadron at the beginning of the sophomore year and continue to take mandatory core courses through graduation. We estimate (1) for academic performance in subsequent years.

Results are shown in Table 5. For comparative purposes, Specification 1 repeats results for freshman year performance. Specifications 2, 3, and 4 reports results for sophomore, junior, and senior performance. Results provide strong evidence that the freshman peer effects persist into follow-on academic performance, but at a diminished rate roughly one-half in magnitude. The freshman peer SAT verbal variable is positive and significant for sophomore (0.176), junior (0.225) and senior (0.198) year academic performance. These results indicate that social network peer effects during the freshman year significantly affect achievement beyond the freshman year. In specifications not shown, we also included freshman GPA as an explanatory variable in the sophomore, junior, and senior grade regressions. In all cases the freshman peer variables were small and statistically insignificant. These results indicate that the freshman squadron peer effects raise an individual's initial GPA and this increase persists throughout a student's career.

5.4 Nonlinearities Across Academic Ability

To this point, all our specifications have been linear-in-mean. However, unless peer effects are nonlinear across individuals (i.e., different across types of students) there would be no social gain to sorting students into peer groups. Taking a high-quality peer from one group and placing her into another group will have an equal and offsetting effect on both groups. To examine potential nonlinearities in the peer effects across academic ability, we compute predicted GPA and estimate separate peer coefficients for each one-third of the (incoming) student academic ability distribution.

Results for this analysis are shown in Table 6 and show a potential nonlinearity in the peer SAT verbal variable. The magnitude of the estimated effect for students in the bottom third of the academic ability distribution (0.565) is larger, although not statistically different, than that estimated for students in the top (0.312) and middle (0.361) third of the distribution. These results indicate that the lowest ability students in the sample may benefit the most from having high quality peers. Placing these low ability students into peer groups with a higher than average SAT verbal score may result in an increased overall average GPA.²⁷

5.5 Falsification Tests

The unique dorm structure at USAFA provides the opportunity to empirically test for false peer effects. All students at USAFA live in one of only two dorm halls. Squadrons 1-21 reside in Vandenberg Hall and squadrons 22-36 reside in Sijan Hall. While all members of a respective squadron are geographically located in the same area of the dorm, squadrons located in the same dorm hall and floor are adjacent to one another with no visible partitions.

To test for the importance of proper identification of the relevant peer group, we are able to construct false peer groups of students whose dorm rooms are located in the same section of the dorm hall, but are not necessarily in the same squadron. We construct these groups using student dorm room assignments at the start of the fall semester. Each dorm room is identified by the hall (Vandenberg or Sijan), floor (2, 3, 5, and 6), section (A to G), and room number. In total, there are 39 identifiable dorm/floor/sections with which we construct false peer groups. These groupings are analogous to hall-floor wings as defined by Foster (2006). During the three years in our sample, 92.3 percent of the hall/floor/sections contain students from different squadrons and the average false peer group is made up of 66.6 percent of members from an individual's actual squadron. We

²⁷We also tested and found little evidence of nonlinearities across gender, race, and athletes.

construct and test for evidence of peer effects using two separate false peer groups: 1) all students within the same hall/floor/section, and 2) freshman students in the same hall/floor/section.

Table 7 presents results for this analysis. Specifications 1 and 2 show results for the first false peer group of all students in the same hall/floor/section with and without controlling for roommates characteristics. Specifications 3 and 4 show results for the second false peer group containing only freshman students in the same dorm/hall/section. In all four specifications, none of the academic peer variables have a statistically significant effect on individual student performance and only the peer fitness variable is positive and significant in Specifications 1 and 2.

Similar to results found by Foster (2006), these results show that geographic proximity of individuals alone does not generate positive peer effects. The false peer groups, on average, contain 67 percent of a student's actual peer group, yet peer effects are virtually undetectable. Again, this highlights the importance of measuring the relevant peer group when estimating peer effects.

6 Conclusion

We examine the random assignment of students to relatively large and tightly controlled social-network peer groups at the United States Air Force Academy for evidence of peer effects in academic performance. The statistical properties of our dataset enable us to identify with great precision the known (exogenous) peer group that an individual spends a majority of his/her time interacting with. Additionally, students in our study have a limited ability to interact with other students outside of their assigned peer group during their freshman year of study. This feature enables us to estimate peer effects that are more comparable to changing the entire cohort of peers.

Our results are significant for several reasons. First, using the broad set of peers an individual spends a majority of his/her time interacting with, we find academic peer effects of much larger magnitude than found in the previous literature. Second, we find the peer effects are largest in the math and science courses and are virtually non-existent in physical education and foreign language courses. These findings suggest the peer effects may be working through study partnerships versus a social norm of effort. Third, we find the freshman peer effects persist at a diminished rate into the sophomore, junior, and senior years. Finally, we find evidence of nonlinearities in the peer effects across student incoming academic ability. These results suggest there may be an opportunity for social gains through sorting individuals into peer groups.

Our results also help explain why many of the previous higher education peer effects studies have found little evidence of positive peer effects in academic performance. We find empirical evidence

that roommates and dorm floors capture only a limited proportion of the total peer influence. As such, we find only moderate evidence of peer influence at the roommate level, as previously found by Sacerdote (2001) and Zimmerman (2003). We also find that geographic proximity of students in dorm halls alone, as in Foster (2006) does not generate measurable peer effects.

References

- Betts, Julian R. and Andrew Zau. 2004. "Peer Groups and Academic Achievement: Panel Evidence from Administrative Data." Unpublished Manuscript.
- Boozer, Michael and Stephen E. Cacciola. 2001. "Inside the 'Black Box' of Project STAR: Estimation of Peer Effects Using Experimental Data." Discussion Paper 832, Yale University.
- Burke, Mary A. and Tim R. Sass. 2008. "Classroom Peer Effects and Student Achievement." Working Paper 08-5, Federal Reserve Bank of Boston. URL <http://ssrn.com/abstract=1260882>.
- Burton, Anthony C., Katherine S. Carson, Susan M. Chilton, and W. George Hutchinson. 2007. "Resolving Questions About Bias in Real and Hypothetical Referenda." *Environmental and Resource Economics* 38 (4):513–525.
- Carrell, Scott E., Frederick V. Malmstrom, and James E. West. 2008. "Peer Effects in Academic Cheating." *Journal of Human Resources* 43 (1):173–207.
- Carrell, Scott E. and James E. West. 2008. "Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors." Working Paper 14081, National Bureau of Economic Research. URL <http://www.nber.org/papers/w14081>.
- Dorans, Neil J. 1999. "Correspondences Between ACT and SAT I Scores." College Board Report 99-1, 99-2, College Entrance Examination Board.
- Falk, Armin and Andrea Ichino. 2006. "Clean Evidence on Peer Effects." *Journal of Labor Economics* 24 (1):39–57.
- Foster, Gigi. 2006. "It's not your peers, and it's not your friends: Some progress toward understanding the educational peer effect mechanism." *Journal of Public Economics* 90 (8-9):1455–1475.
- Guryan, Jonathan, K. Kroft, and M. Notowidigdo. 2007. "Peer Effects in the Workplace: Evidence From Random Assignment in Professional Golf." NBER Working Paper (13422).

- Hanushek, Eric A., John F. Kain, Jacob M. Markman, and Steven G. Rivkin. 2003. "Does Peer Ability Affect Student Achievement?" *Journal of Applied Econometrics* 18 (5):527–544.
- Hoxby, Caroline M. and Gretchen Weingarth. 2006. "Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects." Working Paper.
- Kremer, Michael and Dan M. Levy. 2008. "Peer Effects and Alcohol Use Among College Students." *Journal of Economic Perspectives* 22 (3):189–206.
- Li, Han and Tao Li. 2009. "The Gender Difference of Peer Influence in Higher Education." *Economics of Education Review* 28 (1).
- Lyle, David S. 2007. "Estimating and Interpreting Peer and Role Model Effects from Randomly Assigned Social Groups at West Point." *Review of Economics and Statistics* 89 (2):289–299.
- Manski, Charles F. 1993. "Identification and Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies* 60 (3):531–42.
- Mas, Alexandre and Enrico Moretti. 2009. "Peers at Work." *American Economic Review* 99 (1):112–145.
- Moffitt, Robert A. 2001. "Policy Interventions, Low-Level Equilibria, and Social Interactions." In *Social Dynamics*, edited by Steven Durlauf and H. Peyton Young. MIT Press, 45–82.
- Princeton Review. 2007. "Colleges and Careers." Princeton Review. Accessed on 25 August 2006 at: <http://www.princetonreview.com/college/default.asp>.
- Sacerdote, Bruce L. 2001. "Peer Effects with Random Assignment: Results for Dartmouth Roommates." *Quarterly Journal of Economics* 116 (2):681–704.
- Siegfried, John J. and Michael A. Gleason. 2006. "Academic Roommate Peer Effects." Working Paper.
- Stinebrickner, Ralph and Todd R. Stinebrickner. 2006. "What can be learned about peer effects using college roommates? Evidence from new survey data and students from disadvantaged backgrounds." *Journal of Public Economics* 90 (8-9):1435–54.
- Vigdor, Jacob and Thomas Nechyba. 2007. "Peer Effects in North Carolina Public Schools." In *Schools and the Equal Opportunity Problem*, edited by Ludger Woessmann and Paul E. Peterson. MIT Press, 73–102.

Zimmerman, David J. 2003. "Peer Effects in Academic Outcomes: Evidence From a Natural Experiment." *The Review of Economics and Statistics* 85 (1):9–23.

Table 1: Summary Statistics for Classes of 2005-2007

Variable	Obs	Mean	Std. Dev.	Min	Max
Course Grade (GPA) (freshman fall semester)	19,977	2.91	0.82	-	4.00
SAT Math	3,489	665.47	63.88	440.00	800.00
SAT Verbal	3,489	631.95	67.00	330.00	800.00
Academic Composite	3,488	1,282.41	211.99	623.00	2,067.00
Fitness Score	3,489	459.70	96.88	215.00	745.00
Leadership Composite	3,490	1,724.16	182.42	900.00	2,370.00
Black	3,490	0.05	0.22	0	1
Hispanic	3,490	0.06	0.24	0	1
Asian	3,490	0.05	0.23	0	1
Female	3,490	0.18	0.38	0	1
Recruited Athlete	3,490	0.28	0.45	0	1
Military Preparatory School	3,490	0.21	0.41	0	1
Freshman Roommate SAT Math (mean if two)	2,170	665.95	55.88	460.00	800.00
Freshman Roommate SAT Verbal (mean if two)	2,170	631.11	59.47	350.00	800.00
Freshman Roommate Academic Composite (mean if two)	2,170	1,285.90	188.05	623.00	2,067.00
Freshman Roommate Fitness Score (mean if two)	2,171	458.07	83.81	245.00	735.00
Freshman Roommate Leadership Composite (mean if two)	2,171	1,720.47	160.21	900.00	2,295.00
Peer SAT Math (squadron by class)	108	665.56	12.90	630.00	705.81
Peer SAT Verbal (squadron by class)	108	632.20	11.61	606.97	666.32
Peer Academic Composite (squadron by class)	108	1,282.78	37.70	1,205.41	1,410.58
Peer Fitness Score (squadron by class)	108	459.48	18.12	417.16	507.25
Peer Leadership Composite (squadron by class)	108	1,724.45	31.45	1,625.06	1,795.18

Table 2: Own pre-treatment characteristics regressed on peer pre-treatment characteristics

		Section A: No Controls			Section B: With Randomization Controls		
Class/year		Class of 2005	Class of 2006	Class of 2007	Class of 2005	Class of 2006	Class of 2007
Academic Composite	Freshman	-0.054	0.011	0.049	-0.045	-0.005	0.016
	Roommate	(0.057)	(0.062)	(0.069)	(0.053)	(0.049)	(0.062)
	Freshman	-0.116	0.032	-0.165	-0.032	0.042	-0.130
	Squadron	(0.325)	(0.229)	(0.238)	(0.303)	(0.216)	(0.224)
	Sophomore	-0.117	-0.017	-0.060	-0.172	-0.007	-0.059
	Squadron	(0.288)	(0.166)	(0.240)	(0.262)	(0.163)	(0.236)
SAT Math	Freshman	0.084	-0.034	0.061	0.082	-0.043	0.060
	Roommate	(0.054)	(0.070)	(0.069)	(0.051)	(0.059)	(0.061)
	Freshman	0.255*	-0.055	-0.333	0.328***	0.044	-0.143
	Squadron	(0.146)	(0.364)	(0.325)	(0.120)	(0.318)	(0.261)
	Sophomore	0.120	-0.399	-0.532*	0.152	-0.183	-0.334
	Squadron	(0.206)	(0.319)	(0.281)	(0.155)	(0.293)	(0.244)
SAT Verbal	Freshman	-0.107*	-0.037	-0.032	-0.050	-0.037	-0.030
	Roommate	(0.063)	(0.061)	(0.080)	(0.052)	(0.055)	(0.069)
	Freshman	-0.418	-0.040	-0.578	-0.176	0.069	-0.356
	Squadron	(0.266)	(0.194)	(0.355)	(0.219)	(0.156)	(0.311)
	Sophomore	-0.007	-0.080	-0.712	0.142	0.075	-0.509
	Squadron	(0.309)	(0.312)	(0.449)	(0.238)	(0.266)	(0.377)
Leadership Composite	Freshman	0.014	0.067	-0.020	0.022	0.062	-0.033
	Roommate	(0.057)	(0.079)	(0.055)	(0.056)	(0.081)	(0.061)
	Freshman	-0.574	0.038	0.094	-0.497	0.029	0.119
	Squadron	(0.383)	(0.222)	(0.224)	(0.365)	(0.216)	(0.215)
	Sophomore	0.051	-0.062	-0.124	0.072	-0.042	-0.092
	Squadron	(0.193)	(0.220)	(0.270)	(0.182)	(0.213)	(0.254)
Fitness Score (CFT)	Freshman	0.049	0.002	0.138**	0.031	-0.039	0.145**
	Roommate	(0.069)	(0.042)	(0.057)	(0.068)	(0.045)	(0.055)
	Freshman	-0.110	-0.0004	-0.213	-0.050	-0.045	-0.157
	Squadron	(0.248)	(0.184)	(0.267)	(0.239)	(0.178)	(0.234)
	Sophomore	-0.002	-0.432	-0.289	0.088	-0.391	-0.243
	Squadron	(0.226)	(0.386)	(0.280)	(0.211)	(0.375)	(0.258)

Each coefficient represents a separate regression where the individual (pre-treatment) characteristic is regressed on the peer characteristic. Scrambling controls include indicators for female, black, hispanic, athlete, and preparatory school.

* Significant at the 0.10 level, ** Significant at the 0.05 level, *** Significant at the 0.01 level. Robust standard errors in parentheses are clustered by class by squadron for the squadron level regressions. For individuals with two roommates, the explanatory variables represent the average of the two roommates. For the squadron specifications, the explanatory variables are the average of all classmates in the squadron.

Table 3: Freshman Reduced Form Peer Effects

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Roommate SAT Verbal	0.003 (0.019)				-0.011 (0.021)	-0.008 (0.020)
Roommate SAT Math	-0.005 (0.021)				0.003 (0.022)	-0.003 (0.021)
Roommate Academic Composite	0.0004 (0.005)				0.001 (0.005)	0.001 (0.005)
Roommate Fitness Score	0.022 (0.013)				0.017 (0.014)	0.017 (0.014)
Roommate Leadership Composite	0.012** (0.006)				0.011* (0.006)	0.012** (0.006)
Peer SAT Verbal (other freshmen in squadron)		0.338*** (0.107)		0.412*** (0.107)	0.416*** (0.103)	0.382*** (0.112)
Peer SAT Math (other freshmen in squadron)		-0.119 (0.097)		-0.090 (0.096)	-0.092 (0.106)	
Peer Academic Composite (other freshmen in squadron)		-0.018 (0.032)		-0.023 (0.031)	-0.024 (0.029)	
Peer Fitness Score (other freshmen in squadron)		0.153** (0.064)		0.182*** (0.067)	0.166*** (0.060)	0.145*** (0.065)
Peer Leadership Composite (other freshmen in squadron)		0.024 (0.041)		0.025 (0.042)	0.015 (0.034)	
Sophomore Class SAT Verbal			0.227* (0.131)	0.260** (0.116)	0.259** (0.115)	0.265** (0.117)
Sophomore Class SAT Math			0.053 (0.129)	0.097 (0.128)	0.101 (0.128)	0.088 (0.129)
Sophomore Class Academic Composite			-0.026 (0.030)	-0.006 (0.027)	-0.004 (0.027)	-0.001 (0.027)
Sophomore Class Fitness Score			-0.012 (0.076)	-0.043 (0.067)	-0.041 (0.067)	-0.036 (0.068)
Sophomore Class Leadership Composite			-0.035 (0.039)	-0.015 (0.039)	-0.015 (0.039)	-0.011 (0.039)
Junior Class SAT Verbal			-0.064 (0.116)	0.053 (0.108)	0.047 (0.108)	0.040 (0.110)
Junior Class SAT Math			-0.047 (0.116)	-0.004 (0.111)	-0.038 (0.111)	-0.041 (0.108)
Junior Class Academic Composite			0.005 (0.032)	0.009 (0.032)	0.007 (0.032)	0.003 (0.031)
Junior Class Fitness Score			0.150** (0.075)	0.109 (0.074)	0.106 (0.074)	0.115 (0.073)
Junior Class Leadership Composite			0.049* (0.026)	0.068*** (0.024)	0.068*** (0.024)	0.062*** (0.024)

Table 3: continued

Table 3 (continued)	(1)	(2)	(3)	(4)	(5)	(6)
Senior Class SAT Verbal			0.029 (0.097)	-0.030 (0.104)	-0.034 (0.104)	-0.032 (0.096)
Senior Class SAT Math			0.057 (0.128)	0.033 (0.119)	0.033 (0.119)	0.039 (0.118)
Senior Class Academic Composite			-0.035 (0.027)	-0.053 (0.027)	-0.051 (0.027)	-0.055 (0.027)
Senior Class Fitness Score			0.003 (0.074)	0.002 (0.077)	0.008 (0.078)	0.022 (0.075)
Senior Class Leadership Composite			-0.027 (0.037)	-0.028 (0.035)	-0.026 (0.036)	-0.029 (0.036)
SAT Verbal (own)	0.089*** (0.015)	0.090*** (0.015)	0.087*** (0.016)	0.093*** (0.016)	0.094*** (0.016)	0.094*** (0.016)
SAT Math (own)	0.230*** (0.018)	0.231*** (0.017)	0.233*** (0.018)	0.233*** (0.018)	0.232*** (0.017)	0.232*** (0.017)
Academic Composite (own)	0.110*** (0.004)	0.110*** (0.004)	0.110*** (0.004)	0.109*** (0.004)	0.109*** (0.004)	0.109*** (0.004)
Fitness Score (own)	0.052*** (0.009)	0.052*** (0.009)	0.052*** (0.009)	0.055*** (0.009)	0.054*** (0.009)	0.054*** (0.009)
Leadership Composite (own)	0.004 (0.004)	0.004 (0.004)	0.004 (0.004)	0.004 (0.004)	0.003 (0.004)	0.003 (0.004)
Observations	19,966	19,966	19,966	19,966	19,966	19,966
R ²	0.3468	0.3478	0.3482	0.3504	0.3509	0.3507
F-statistic p-value (5, 107): roommate variables	0.167				0.329	0.321
F-statistic p-value (5, 107): peer variables		0.021		0.004	0.008	0.001
F-statistic p-value (15, 107): upperclass variables			0.006	0.001	0.001	0.001
F-statistic p-value (20, 107): peer and upperclass variables				0.001	0.001	0.0004
F-statistic p-value (25, 107): roommate, peer, and upperclass					0.0003	0.0001
Control Variables	course by section fixed effects	course by section fixed effects	course by section fixed effects	course by section fixed effects	course by section fixed effects	course by section fixed effects

The dependent variable in each specification is the course grade in the freshman fall semester. Each observation is weighted by the number of course credit hours. Missing roommate pre-treatment data were imputed using the cohort mean of each variable. * Significant at the 0.10 level, ** Significant at the 0.05 level, *** Significant at the 0.01 level. Robust standard errors in parentheses are clustered by class by squadron. All specifications include individual-level controls for students who are black, Hispanic, Asian, female, recruited athlete, attended a preparatory school, and missing roommate data.

Table 4: Peer Effects by Course Type

	All Courses	Math & Science	Humanities and Social Science	Foreign Language	Physical Education	Military Studies
	(1)	(2)	(3)	(4)	(5)	(6)
Peer SAT Verbal (other freshmen in squadron)	0.416*** (0.103)	0.509*** (0.132)	0.405*** (0.154)	0.163 (0.201)	0.132 (0.091)	0.146 (0.101)
Peer SAT Math (other freshmen in squadron)	-0.092 (0.106)	-0.098 (0.122)	-0.078 (0.151)	-0.026 (0.159)	-0.134* (0.072)	-0.011 (0.100)
Peer Academic Composite (other freshmen in squadron)	-0.024 (0.029)	-0.042 (0.036)	-0.014 (0.037)	0.018 (0.056)	-0.023 (0.024)	-0.015 (0.030)
Peer Fitness Score (other freshmen in squadron)	0.166*** (0.060)	0.168** (0.074)	0.204** (0.087)	-0.138 (0.118)	0.116** (0.053)	0.151** (0.064)
Peer Leadership Composite (other freshmen in squadron)	0.015 (0.034)	0.015 (0.045)	-0.012 (0.046)	0.072 (0.058)	0.043* (0.025)	0.073* (0.038)
Observations	19,966	9,313	3,946	1,425	3,333	3,120
R ²	0.3509	0.3652	0.3337	0.3218	0.5705	0.3552
Control Variables	roommate and upperclass peer variables, course by section fixed effects, year fixed effects	roommate and upperclass peer variables, course by section fixed effects, year fixed effects	roommate and upperclass peer variables, course by section fixed effects, year fixed effects	roommate and upperclass peer variables, course by section fixed effects, year fixed effects	roommate and upperclass peer variables, course by section fixed effects, year fixed effects	roommate and upperclass peer variables, course by section fixed effects, year fixed effects

The dependent variable in each specification is the course grade in the freshman fall semester. Each observation is weighted by the number of course credit hours. Missing roommate pre-treatment data were imputed using the cohort mean of each variable. * Significant at the 0.10 level, ** Significant at the 0.05 level, *** Significant at the 0.01 level. Robust standard errors in parentheses are clustered by class by squadron. All specifications include individual-level controls for students who are black, Hispanic, Asian, female, recruited athlete, attended a preparatory school, and missing roommate data.

Table 5: Persistence in the Freshman Peer Group Effects

	Freshman Grades	Sophomore Grades	Junior Grades	Senior Grades
	(1)	(2)	(3)	(4)
Peer SAT Verbal (other freshmen in squadron)	0.416*** (0.103)	0.176** (0.082)	0.225** (0.091)	0.198** (0.071)
Peer SAT Math (other freshmen in squadron)	-0.092 (0.106)	-0.009 (0.086)	0.088 (0.086)	0.001 (0.061)
Peer Academic Composite (other freshmen in squadron)	-0.024 (0.029)	0.004 (0.024)	-0.005 (0.027)	-0.039** (0.018)
Peer Fitness Score (other freshmen in squadron)	0.166*** (0.060)	-0.031 (0.057)	0.031 (0.059)	-0.095** (0.044)
Peer Leadership Composite (other freshmen in squadron)	0.015 (0.034)	0.018 (0.037)	0.013 (0.037)	0.026 (0.027)
Observations	19,966	23,200	20,297	17,825
R ²	0.3509	0.3254	0.3501	0.3801
Control Variables	roommate and upperclass peer variables, course by section fixed effects, year fixed effects	roommate and upperclass peer variables, course by section fixed effects, year fixed effects	roommate and upperclass peer variables, course by section fixed effects, year fixed effects	roommate and upperclass peer variables, course by section fixed effects, year fixed effects

The dependent variable in each specification is the course grade in the fall semester. Each observation is weighted by the number of course credit hours. Missing roommate pre-treatment data were imputed using the cohort mean of each variable. * Significant at the 0.10 level, ** Significant at the 0.05 level, *** Significant at the 0.01 level. Robust standard errors in parentheses are clustered by class by squadron. All specifications include individual-level controls for students who are black, Hispanic, Asian, female, recruited athlete, attended a preparatory school, and missing roommate data.

Table 6: Nonlinearities in the Freshman Peer Group Effects

Predicted GPA using pre-treatment characteristics	<u>Bottom</u>	<u>Middle</u>	<u>Top</u>
Peer SAT Verbal (other freshmen in squadron)	0.565*** (0.159)	0.361** (0.154)	0.312** (0.132)
Peer SAT Math (other freshmen in squadron)	-0.198 (0.146)	0.006 (0.140)	-0.056 (0.112)
Peer Academic Composite (other freshmen in squadron)	-0.067 (0.045)	-0.042 (0.041)	0.022 (0.036)
Peer Fitness Score (other freshmen in squadron)	0.154*** (0.107)	0.184*** (0.091)	0.169*** (0.074)
Peer Leadership Composite (other freshmen in squadron)	0.057 (0.057)	0.006 (0.050)	0.007 (0.046)
Observations		19,966	
R ²		0.3526	
Control Variables	roommate and upperclass peer variables, course by section fixed effects, year fixed effects		

The dependent variable in each specification is the course grade in the fall semester. Each observation is weighted by the number of course credit hours. Missing roommate pre-treatment data were imputed using the cohort mean of each variable. Separate coefficients are estimated for students in each third of the incoming academic ability distribution. * Significant at the 0.10 level, ** Significant at the 0.05 level, *** Significant at the 0.01 level. Robust standard errors in parentheses are clustered by class by squadron. All specifications include individual-level controls for students who are black, Hispanic, Asian, female, recruited athlete, attended a preparatory school, and missing roommate data.

Table 7: Peer Falsification Tests

	Section A: False Peer 1		Section B: False Peer 2	
	(1)	(2)	(3)	(4)
Peer SAT Verbal	-0.062 (0.186)	-0.034 (0.184)	0.014 (0.086)	0.025 (0.086)
Peer SAT Math	-0.160 (0.171)	-0.156 (0.167)	-0.027 (0.082)	-0.034 (0.079)
Peer Academic Composite	-0.020 (0.060)	-0.019 (0.061)	-0.014 (0.028)	-0.012 (0.028)
Peer Fitness Score	0.228** (0.115)	0.241** (0.115)	-0.019 (0.059)	0.001 (0.058)
Peer Leadership Composite	0.004 (0.073)	0.014 (0.074)	0.008 (0.033)	0.019 (0.033)
Observations	19,740	19,740	19,740	19,740
R ²	0.3509	0.3502	0.3503	0.3496
F-statistic p-value (5, 107): peer variables	0.353	0.329	0.990	0.981
Control Variables	roommates peer variables, upperclass peer variables, year fixed effects	upperclass peer variables, year fixed effects	roommates peer variables, upperclass peer variables, year fixed effects	upperclass peer variables, year fixed effects

The dependent variable in each specification is the course grade in the freshman fall semester.

Missing roommate pre-treatment data were imputed using the cohort mean of each variable. Peer coefficients represent artificially constructed peer groups. False Peer 1 represents all students in the same dorm section and False Peer 2 represents all freshman students in the same dorm section. * Significant at the 0.10 level, ** Significant at the 0.05 level, *** Significant at the 0.01 level. Robust standard errors in parentheses are clustered by class by false peer group. All specifications include individual-level controls for students who are black, Hispanic, Asian, female, recruited athlete, attended a preparatory school, and missing roommate data. There are 123 missing students from the full sample due to missing room numbers.