NBER WORKING PAPER SERIES

DOES YOUR COHORT MATTER? MEASURING PEER EFFECTS IN COLLEGE ACHIEVEMENT

Scott E. Carrell
Richard L. Fullerton
James E. West

Working Paper 14032
http://www.nber.org/papers/w14032

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2008

Special thanks go to: William Carpenter, Rolland Stoneman, Kathleen O'Donnell, JauTsau, 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. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Does Your Cohort Matter? Measuring Peer Effects in College Achievement
Scott E. Carrell, Richard L. Fullerton, and James E. West
NBER Working Paper No. 14032
May 2008
JEL No. I20

ABSTRACT

To estimate peer effects in college achievement we exploit a unique dataset in which individuals have been 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 find academic peer effects of much larger magnitude than found in previous studies that have measured peer effects among roommates alone. We find the peer effects persist at a diminishing rate into the sophomore, junior, and senior years, indicating social network peer effects may have long lasting effects on academic achievement. Our findings also suggest that peer effects may be working through study partnerships versus operating through establishment of a social norm of effort.

Scott E. Carrell
Department of Economics
University of California, Davis
One Shields Avenue
Davis, CA 95616
and NBER
secarrell@ucdavis.edu

James E. West
Department of Economics and Geosciences
U.S. Air Force Acdemy
2354 Fairchild Dr. #6K100
USAF Academy, CO 80840
Jim.West@usafa.edu

Richard L. Fullerton
Department of Economics and Geosciences
U.S. Air Force Academy #6K100
2354 Fairchild Hall
USAF Academy, CO 80840
richard.fullerton@usafa.edu

1 Introduction

Students spend vast resources on the college admissions process and will often pay thousands of dollars in additional tuition to attend a more prestigious college or university. But what does this really buy? Researchers have generally found there is a labor market premium associated with attending a more selective college.[1] Attending a more prestigious college presumably buys individuals three primary things: 1) access to greater resources; 2) better peers and social networks; and, 3) a potential positive signal to the labor market upon graduation. With regard to peers, relatively little is known about how the quality of ones cohort of postsecondary peers affects individual outcomes. That is, how would student outcomes differ by having a cohort of higher quality peers, as one would get at a more selective university, 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.[2] Results from these studies have found only mixed evidence regarding the existence of positive peer effects in academic performance. [3] A major drawback of these studies however, is that roommates are generally only a small subset of an individuals actual peer group.[4] 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 individuals 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 have been exogenously assigned to peer groups of about 30 students, with whom they are required to spend the majority of their

[1]See: Hoeksta (Forthcoming), Dale and Krueger (2002), Brewer, Eide, and Ehrenberg (1996), Brewer, Eide, and Ehrenberg (1999), Eide, Brewer, and Ehrenberg (1998), Behrman, Rosenzweig, and Taubman (1996), Jones (1990), Mueller (1988) and Loury and Garman (1995).

[2]The one notable exception is Lyle (2007) who examines peer effects using data from the US Military Academy (USMA).

[3] To date there has been little evidence of large positive peer effects in academic performance. For example, 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 Barea College. Also see: Hoxby and Weingarth (2006), Siegfried and Gleason (2006), Li and Li (Forthcoming) and Kremer and Levy (2003).

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

time interacting. Conditional on a few demographic characteristics[5], 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.[6] This feature enables us to estimate peer effects that are more comparable to changing the entire cohort of peers. Additionally, students are 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, when we use the broad peer group to which the individual is assigned as the definition of the peer group, we find academic peer effects that are much larger in magnitude than the previous literature. For freshman students, a 100-point increase in the peer group average SAT verbal score increases individual GPA by 0.45 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 operating through establishment of a social norm of effort.[7] Third, we examine the persistence of the freshman peer effects over the entire four-year academic career. Results show the freshman peer group effects persist at a diminishing rate into the sophomore, junior, and senior years, indicating that social network peer effects may have long lasting effects on academic achievement.

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

Our data come from the United States Air Force Academy (USAFA), a small undergraduate college with an approximate enrollment of 4,200 students. We recognize that questions could be

[5]Females, minorities, athletes, and students who attended a preparatory school are randomly sorted into peer groups first, to ensure diversity across groups.

[6]The two major exceptions are during academic lectures and students who participate in intercollegiate athletes.

[7] Not all students are required to take a foreign language and students are spread across foreign language courses in Spanish, French, German, Russian, Chinese, Japanese, and Arabic. Given this, the likelihood of finding a suitable study partner within a given squadron is much smaller than for other freshman courses, which have near universal common enrollment.

raised about the generalizeability of our findings given USAFA students are a subset of traditional college students. However, our study would not be possible without the random assignment of students into clearly identified, non-overlapping peer groups which can be measured and tracked. Despite the military academy setting, much about USAFA is comparable to broader academia. USAFA faculty have earned their graduate degrees from a broad sample of high quality programs in their respective fields, as would be found in a comparable undergraduate liberal arts college. While the Air Force Academy student body includes a 17 percent of students whose parents were in the military themselves, the rest of the students are 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, Carson, Chilton, and Hutchinson (2007) find the behavior of USAFA students and students at Queens University, Belfast to be statistically indistinguishable.

We also recognize that 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.

The remainder of the paper unfolds as follows. Section II reviews the challenges in measuring peer effects and describes the evaluation strategy used in this paper. Section III describes the peer group structure in our study. Section IV presents the data and its relevance for the measurement of peer effects. Section V presents the methods and results. Section VI concludes.

2 Measuring Peer Effects

Manski (1993) distinguishes three types of peer influence: 1) endogenous effects, 2) exogenous effects, and 3) correlated effects. Endogenous effects occur when individual behavior varies with the behavior of the group. Exogenous or contextual effects occur when individual behavior varies with the pre-treatment group characteristics. Finally, correlated effects are those driven by common treatments. For example, in college academic achievement measured by a GPA, the endogenous effects are those that vary with the average GPA performance of the peer group. Exogenous effects are those that vary with the socio-economic status or the high school performance of the peer

group. Correlated effects are those that are driven by common shocks, such as teachers or dorm room quality.

Measuring the importance of each of these effects is difficult for two main reasons. First, it is difficult to separate out the individual and group influence on one another (Vigdor and Nechyba 2004). This problem is often referred to as the endogeneity problem (Moffitt 2001, Sacerdote 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 with respect to the stochastic error component of the dependent variable. A more recent strategy in the education peer effects literature has used previous peer achievement as an instrument for current achievement (Betts and Zau 2004, Burke and Sass 2004, Hanushek, Kain, Markman, and Rivkin 2003, Vigdor and Nechyba 2004).

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 the variation across classrooms or cohorts within a school.[8] This has typically been accomplished using large administrative panel data sets while employing a series of fixed effects models. 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.[9]

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 provides several new insights compared to the previous literature.[10] 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, our vast amount of exogenous pre-treatment data allows us to correct for

[8] See: Carrell, Malmstrom, and West (2008),Hoxby and Weingarth (2006),Vigdor and Nechyba (2004), Betts and Zau (2004), Burke and Sass (2004), Hanushek, Kain, Markman, and Rivkin (2003), among others.

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

[10] The one paper with a similar identification strategy to ours is Lyle (2007) who primarily examines the effects of common shocks on estimating contemporaneous peer effects using data from the US Military Academy (USMA). However, as Lyle notes that USMA peer groups are constructed to have similar distributions of academic aptitude. This leveling of academic aptitude by peer group results in a reduction of exogenous variation in the peer pre-treatment variables. USAFA, by contrast, has random assignment into peer groups with respect to academic aptitude. The effect of this is that the exogenous variable academic aptitude as measured by peer SAT scores has 49−percent more exogenous variation in our data compared to Lyle (2007).

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 courses to estimate the peer effects across different types of courses free from selection bias into or out of course or section. Finally, reassignment to a new squadron peer group in the sophomore year allows us to test for the persistence in the peer effects over time.

We estimate peer effects using reduced form equations, where we regress individual outcomes on pre-treatment variables to avoid simultaneous equation bias or the reflection problem. We use a variety of own, roommate, peer (other freshmen in squadron), and upperclassmen pre-treatment variables. Freshman GPA is presumed to be exogenous with respect to such variables as SAT scores (both math and verbal), academic composite (to include high school GPA, class rank, quality of school, size of school), fitness scores, and leadership composite scores required for entry to USAFA. Our specification uses the linear-in-means model common to the peer effects literature. While we recognize the potential policy limitations of linear-in-means models (Hoxby and Weingarth 2006, Weinberg 2005), we use it to identify the average peer effect across our entire population.

In general, we find strong, robust peer effects of larger magnitude than those found in previous studies. We credit this to randomized peer group formation, the copious amounts of data that USAFA keeps on all students, and the nature of the squadron structure, which allows us to cleanly identify the group of possible peers for freshman students.

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 $4,200$ students. There are 32 majors offered including 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. In addition, applicants must receive a nomination from a legal nominating authority including Members of Congress, the Vice President, or President of the United States, and other related sources. All students attending the Air Force Academy receive 100% scholarship to cover their tuition, room, and board. Additionally, each student receives a monthly stipend of $845 to cover books, uniforms, computer, and other

living expenses. All students are required to graduate within four years[11] and serve a five-year commitment as a commissioned officer in the United States Air Force following graduation.

Students are grouped into one of 36 peer groups, called squadrons, with each group comprised of approximately 120 students (freshman through seniors). Students of a squadron live in adjacent dorm rooms, dine together, compete in intramural sports together, perform military training together, and study together. 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.[12]

A significant amount of social, academic, athletic, and leadership 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. In theory, any member of the squadron could potentially help a freshman student with his/her coursework. As freshman students are junior, 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 made easy by the way the Academy splits students between squadrons. Upon admission, conditional on a few demographic characteristics, freshman students are randomly assigned to a squadron, and randomly assigned to a roommate within their squadron. This structure creates a natural experiment for estimating peer influence. The overwhelming majority of entering students do not know anybody currently enrolled at USAFA. Sibling students are deliberately separated. 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.

As freshman roommate and squadron assignments are accomplished without **any** input from freshman students, self-selection into squadrons is not a concern. 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

[11] Special exceptions are given for religious missions, medical "set-backs", and other instances beyond the control of the individual.

[12] Students are intermixed during academic classes and can meet with students from other squadrons at the library, gym, church, and what would be considered the student union. Additionally, freshman students who are on intercollegiate athletic teams or participate in club sports are intermixed with students from other squadrons during practice times and on team trips.

the effect the institution is trying to achieve in bypassing student preferences (and, fortunately 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.

Students are re-assigned to a new squadron at the start of their sophomore year and remain in that squadron for the next three years. This feature of the USAFA system enables us to test for the persistence of the freshman peer effects on performance throughout students entire four-year academic history.

4 Data

4.1 The Dataset

Data on students pre-Academy characteristics and on their performance while at the Academy were provided by USAFA Institutional Research and Assessment and de-identified by the USAFA Institutional Review Board. A complete list of summary statistics is provided in Table 1.[13]

Our dataset includes all students in the graduating classes of 2000 through 2007. 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. Seven-percent of students at USAFA have a parent who graduated from a service academy and 17-percent have a parent who served in the military.

Pre-Academy (pre-treatment) data includes whether students were recruited as athletes, whether they attended a military preparatory school, and measures of their academic, athletic and leadership aptitude. Pre-treatment academic aptitude is measured through *SAT verbal* and *SAT math* scores and an *academic composite* computed by the USAFA admissions office, which is a weighted average of an individuals high school GPA, class rank, and the quality of the high school attended. The sample mean SAT math, SAT verbal, and academic composite are 665, 643, and $1,282$ with respective standard deviations of 64, 67, and 212. The measure of pre-treatment athletic aptitude consists of a score on a fitness test(*fitness score*), required by all applicants prior to entrance.[14] The sample mean fitness score is 460 with a standard deviation of 97. The measure of pre-treatment

[13] As fully discussed in the next section, due to concerns with potential non-random placement of students into squadrons prior to the class of 2005, the summary statistics provided only include the graduating classes of 2005-2007.

[14] The fitness score measures timed scores in pull-ups, sit-ups, push-ups and a 600-yard shuttle run, in addition to a standing long jump and a basketball throw.

leadership aptitude is a *leadership composite* computed by the USAFA admissions office, which is a weighted average of high school and community activities (e.g., student council offices, Eagle Scout, captain of sports team, etc.). The sample mean leadership composite is $1,724$ with a standard deviation of 183.

Our outcome performance data contains each individuals freshman through senior academic performance as measured by a grade point average (GPA) computed on a zero to 4.0 scale. Grades are determined on an A, A-, B+, B $\cdots$ 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. GPA is 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).[15] 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 the same grading scale for all students in the course, regardless of instructor. This institutional characteristic assures there is no self-selection of students into courses or towards certain professors.

The absence of self-selection into courses or to professors allows us to rule out potential mechanisms driving our peer effects results. First, we know peers influencing the choice of courses, professors, or academic major do not drive the results. Second, as students from squadrons are randomly mixed across classrooms, our peer effects are not driven through classroom peer interactions or common shocks within the classroom.

4.2 Are Peer Group and Freshman Roommate Assignments Truly Random?

We obtained the algorithm that placed students into squadrons for the classes of 2005 through 2007 from the USAFA Admissions Office.[16] The algorithm prevents siblings as well as students within

[15] Carrell 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.

[16] We have been unable to obtain the algorithm that placed students into squadrons prior to the class of 2005. However, we were informed that the algorithm was rewritten starting in 2000, just prior to the class of 2004 entering, when the admissions office migrated from a Unisys to an Oracle-based system. The timing of the migration from Unisys to Oracle is consistent with the observed changes in squadron selection bias between the classes of 2004 and 2005. Officials in the USAFA Admissions Office acknowledge the possibility of minor changes being implemented to

the same graduating class or with the same last name from being placed in the same squadron. Additionally, females, minorities, athletes, and students who attended a military preparatory school are randomly sorted into squadrons first, to ensure diversity across squadrons. The rest of the students, however, are then randomly assigned to a squadron. Of prime importance to our study is that students are indeed 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*).[17]

Freshman squadron placements were unavailable for the graduating classes of 2000, 2001, and 2003; therefore, results for these classes only include sophomore squadron assignments. 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 2/3 of freshman students as roommates. We considered individuals as roommates if students were issued a key to the same room for a minimum of 2 overlapping months.

The negative and highly significant coefficients on the freshman squadron peer academic and peer athletic composite variables for the classes of 2002 and 2004 indicates a negative selection effect on freshman squadron placements during these years (Table 2). These results suggest that USAFA personnel may have sorted students into squadrons based on pre-treatment characteristics during these years with the intention of balancing each squadrons overall academic and/or athletic ability. Sophomore squadron placements appear to have the same negative selection for the class of 2003 (Table 2). This negative selection, which reduces or eliminates exogenous variation in pre-treatment characteristics across groups, would lead to negatively biased peer effects estimates.[18]

There appears to be little evidence of squadron selection effects in the data for the classes of 2005 through 2007, with all but one selection coefficient statistically insignificant at the 0.05-level

the sorting algorithm when it was migrated from Unisys to Oracle, and that such changes could have been implemented without office memoranda documenting such a change.

[17] Squadron size in our sample averages 32.7. If students are randomly placed into squadrons, 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. For example, the standard deviation of peer group SAT verbal score is 11.4 where population standard deviation is $67.0/\sqrt{32.7} = 11.7$. If instead students had been sorted into squadrons so as to minimize squadron variance of academic ability, we would expect measured squadron variation to be less.

[18] Lyle (2007) notes, "It is possible that the scrambling process reduces the variation in average pretreatment ability measures to the extent that no effect is identifiable. "

(Table 2).[19] At the roommate level, the one exception is a positive and significant coefficient on the roommate fitness score for the class of 2007, indicating a potential positive selection of roommates on athletic ability. However, this positive coefficient diminishes and is statistically insignificant when including a squadron fixed-effect, indicating that within squadrons, where roommates are assigned, there appears to be no positive selection.

Based on these findings and the absence of specific information regarding the squadron assignment process prior to the class of 2005, we restrict our sample to the classes 2005 through 2007. By doing so, we ensure that there is adequate exogenous variation in the mean pre-treatment characteristics across peer groups.

5 Methods and Results

We analyze the peer effects using the traditional 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:

$$GPA_{isc} = \phi_0 + \phi_1 X^r_{isc} + \phi_2 \frac{\sum\limits_{k \neq i} X_{ksc}}{n_{sc} - 1} + \beta X_{isc} + \epsilon_{isc} \tag{1}$$

where GPA_{isc} is the freshman fall semester GPA for individual i in squadron s, graduating class c. X^r_{isc} are the pre-treatment characteristics of individual is roommate[20]and $\frac{\sum\limits_{k \neq i} X_{ksc}}{n_{sc}-1}$ are the average pre-treatment characteristics of all other peers in squadron s except individual i. X_{isc} 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, and whether they attended a military preparatory school. ϵ_{isc} is the error term. We include graduating class fixed effects to control for unobserved mean differences across years in GPA. 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.

[19] At the 0.10-level, SAT math is positive and significant for the class of 2005 and negative and significant at the 0.10-level for the class of 2007. However, with 45 selection regressions and random sampling, one would expect at least 4 coefficients to be significant at the 0.10-level. Additionally, there is no evidence of selection bias on academic ability when performing these same regressions using the USAFA admission offices total academic composite, which combines SAT math, SAT verbal, high school GPA, class rank, and the quality of high school attended.

[20] Average GPA is used for individual with two roommates.

5.1 Main Results

We estimate various specifications of equation (1) using ordinary least squares (OLS) for freshman academic performance, with results shown in Table 3.[21] For Specification 1, we estimate the peer influence at the roommate levelusing the full array of roommate-level academic, athletic, and leadership pre-treatment measures.[22] We find insignificant coefficients for the roommate SAT verbal, SAT math, academic composite and fitness score variables; however, the coefficient on the roommate leadership composite is positive and significant (0.013) at the 0.05-level. The effect is relatively small; the model predicts a one-standard deviation increase in the roommate leadership composite results in an increased freshman fall semester GPA of 0.02 grade points. The F-statistic (1.53) for the five roommate variables is statistically insignificant, indicating that roommate pre-treatment characteristics alone do not provide statistically significant explanatory power. Own SAT verbal (0.059), SAT math (0.240), academic composite (0.109) and fitness score (0.045) 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 is peers (other freshmen) in squadron s. Of the five peer variables estimated, two coefficients are statistically significant, peer SAT verbal (0.348) and peer fitness score (0.139). The F-statistic (2.32) on the five peer variables is significant at the 0.05-level providing evidence that this broader peer group plays a more important role than that of roommates. Compared to previous studies, the magnitude of peer SAT verbal is quite large, and 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 squadronto measure the leadership effects from the upperclassmen within the squadron. Of the 15 upper class variables estimated, only the junior class leadership composite (0.059) is individually significant; however, all fifteen variables are jointly significant at the 0.05-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 the peer pre-treatment characteristics are jointly significant at the 0.01-level and the upper class characteristics are jointly significant at the 0.05-level.

[21] SAT scores, academic composite, leadership composite and fitness scores have all been divided by 100 prior to estimating the regressions.

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

In Specification 5 we estimate the model using pre-treatment characteristics of individual is 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 (0.013), 2) peer SAT verbal (0.448), 3) peer fitness score (0.153), 4) sophomore class SAT verbal (0.284), and 5) junior class leadership composite (0.104). The positive results for the roommate leadership composite, peer SAT verbal, and peer fitness test variables provide evidence of positive peer influence and the positive results for the sophomore class SAT verbal and junior class leadership composite variables provide evidence of positive leadership effects within the squadron. All 25 roommate, peer, and upper class pre-treatment characteristics are jointly significant at the 0.01-level (F-statistic = 2.73), providing evidence that peers and leaders play a significant role in the academic performance of the freshman within the squadron.

The previous results provide strong evidence of positive social spillovers in academic performance.[23] 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 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. The model estimates that a one-standard deviation increase in the peer SAT verbal score results in an increased own GPA of 0.052 grade points (one-twelfth of a standard deviation). In terms of standard deviations, this effect size is nearly 2.5 times greater in magnitude compared to that found by Zimmerman (2003) for roommates at Williams College.[24]

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 prone to peer influence, we would also expect to see larger peer effects at the roommate-level compared to previous studies. Thus, the absence of large effects at the roommate-level indicates the institutional setting at USAFA is not solely driving the results. We next explore the importance of properly identifying the relevant peer group.

[23] For brevity we do not show the reduced form estimates on athletic performance. In these specifications, we find only one positive and statistically significant effect (junior class leadership composite). However, the peer and upper class pre-treatment characteristics are jointly significant at the 0.05-level.

[24] Zimmerman (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.

5.2 Falsification Tests

The unique dorm structure at USAFA provides the opportunity to empirically test for false peer effects. All 4,200 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 individuals actual squadron. We construct and test for 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 4 presents results for this analysis. Specifications 1 and 2 show results for the first false peer of all students in the same hall/floor/section with and without controlling for roommates. 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 Specification 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 and highlight the importance of measuring the relevant peer group when estimating peer effects. The false peer groups, on average, contain 67 percent of a students actual peer group, yet peer effects are virtually undetectable.

5.3 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. We estimate these effects as in equation (1) while adding a course-by-section fixed effect. The course-by-section fixed effects control for all classroom peer effects and differences in instructor quality.

Table 5 presents results for this analysis. Specification 1 shows results for math and science courses. The magnitude of the coefficient for the peer SAT verbal variable (0.672) is very large and highly significant. The magnitude of the effect is roughly 50 percent larger than we previously measured using fall semester GPA. The model estimates that a 1-standard deviation increase in the peer SAT verbal variable increases math and science performance by 0.08 grade points. Specification 2 shows results for humanities and social science courses. Again the peer SAT verbal variable is positive and statistically significant (0.435), with the magnitude of the effect smaller than that found in the math and science courses.

Specifications 3 and 4 present results for foreign language and physical education courses. In both specifications there is almost no evidence of a peer effect. The foreign language results are not surprising as not all students are required to take a foreign language. Additionally, the students who take a foreign language are spread across taking Spanish, French, German, Russian, Chinese, Japanese, and Arabic. Thus, the opportunity for peer interaction within a squadron is very limited. The results for the physical education courses are also not surprising and somewhat reassuring as there is virtually no work outside of class for physical education courses.

Finally, Specification 5 shows results for the military studies courses. The peer SAT verbal variable is positive and statistically significant (0.289) as is the peer fitness variable (0.154) and the peer leadership variable (0.122).

The preceding results show the peer effects are largest in the math and science courses and 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 because physical education and foreign language courses have the least opportunities for interaction among students. We next explore the persistence of the freshman squadron peer effects into the sophomore, junior, and senior year.

5.4 Persistence of the Effects

With evidence of positive peer effects in freshman academic performance, we next examine the persistence of freshman peer effects in performance in follow-on years. It is possible to statistically separate 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. Additionally, students at USAFA take 30+ core courses. We estimate equation (1) for follow-on year(s) grade performance while including current squadron-by-year fixed effects and course-by-section fixed effects. The squadron-by-year fixed effects control for all contemporaneous peer interactions and therefore isolates the freshman peer effects on follow-on academic performance. The course-by-section fixed effects control for all classroom peer effects and differences in instructor quality.

Results are shown in Table 6. For comparative purposes, Specification 1 presents results for freshman year performance using individual-level grades for the entire academic year. Results show that the peer SAT verbal variable (0.519) and the peer fitness score variable (0.126) are positive and highly significant. Additionally, the peer academic composite variable is negative, small (-0.067) and marginally significant. Specifications 2, 3, and 4 show 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 diminishing rate. For specification 2, the peer SAT verbal variable (0.023) is positive and significant and is roughly one-half in magnitude of that found in the freshman year. The peer SAT verbal variable continues to be positive and significant in the junior (0.020) and senior (0.176) years, with slight decreases in magnitude. Additionally, the peer fitness variables are statistically insignificant in all follow-on year specifications.

The results for the peer SAT verbal variables indicate the freshman squadron peer effects persist at a diminishing rate into academic performance in the sophomore, junior, and senior years after students are reassigned to a new squadron. These results indicate that social network peer effects may have long lasting effects on academic achievement. 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 individuals initial GPA and this increase persists throughout a students career.[25]

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

[25]In specifications not shown we also computed a model where the dependent variable was the change in GPA from the freshman to sophomore year. The coefficient for the peer SAT verbal score was negative. This result confirms the previous results and indicates that the peer effect raises the level of an individuals GPA and this increased level persists, but at a diminishing rate.

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 their time interacting with, we find academic peer effects of much larger magnitude than found in the previous literature. For freshman students a 100-point increase in the peer group average SAT verbal score increases individual GPA by 0.45 grade points. Second, using course-level data 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 as physical education and foreign language courses have the least opportunities for interaction among students.[26] Third, we examine the persistence of the peer effects over the entire four-year academic career. Results show the freshman peer group effects persist at a diminishing rate into the sophomore, junior, and senior years, indicating that social network peer effects may have long lasting effects on academic achievement.

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

Behrman, J. R., M. R. Rosenzweig, and P. Taubman (1996): "College Choice and Wages: Estimates Using Data on Female Twins," *Review of Economics and Statistics*, LXXXVII, 672–685.

[26] Not all students are required to take a foreign language and students are spread across foreign language courses in Spanish, French, German, Russian, Chinese, Japanese, and Arabic. Given this, the likelihood of finding a suitable study partner within a given squadron is much smaller than for other freshman courses, which have near universal common enrollment.

Betts, J. R., and A. Zau (2004): "Peer Groups and Academic Achievement: Panel Evidence from Administrative Data," Unpublished Manuscript.

Boozer, M., and S. E. Cacciola (2001): "Inside the 'Black Box' of Project STAR: Estimation of Peer Effects Using Experimental Data," Unpublished Manuscript.

Brewer, D., E. Eide, and R. Ehrenberg (1996): "Does It Pay to Attend an Elite Private College? Evidence from the Senior High School Class of 1980," *Research in Labor Economics*, 15, 239–271.

——— (1999): "Does It Pay to Attend an Elite Private College? Evidence on the Effects of College Type on Earnings," *Journal of Human Resources*, 34, 104–123.

Burke, M. A., and T. R. Sass (2004): "Classroom Peer Effects and student Achievement," in *Presented at the American Economic Association Annual Meeting, January 2005*.

Burton, A. C., K. S. Carson, S. M. Chilton, and W. G. Hutchinson (2007): "Resolving Questions About Bias in Real and Hypothetical Referenda," *Environmental and Resource Economics*, 38(4), 513–525.

Carrell, S. E., F. V. Malmstrom, and J. E. West (2008): "Peer Effects in Academic Cheating," *Journal of Human Resources*, 43(1), 173–207.

Carrell, S. E., and J. E. West (2008): "Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors," Unpublished Manuscript.

Dale, S. B., and A. B. Krueger (2002): "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables," *Quarterly Journal of Economics*, 17(4), 1491–1527.

Dorans, N. J. (1999): "Correspondences Between ACT and SAT I Scores," College Board Report 99-1, 99-2, College Entrance Examination Board.

Eide, E., D. Brewer, and R. Ehrenberg (1998): "Does It Pay to Attend an Elite Private College? Evidence on the Effects of Undergraduate College Quality on Graduate School Attendance," *Economics of Education Review*, 17, 371–376.

Foster, G. (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.

Hanushek, E. A., J. F. Kain, J. M. Markman, and S. G. Rivkin (2003): "Does Peer Ability Affect Student Achievement?," *Journal of Applied Econometrics*, 18(5), 527–544.

Hoeksta, M. L. (Forthcoming): "The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach," *Review of Economics and Statistics*.

Hoxby, C. M., and G. Weingarth (2006): "Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects," Working Paper.

Jones, E. (1990): "College Grades and Labor Market Outcomes," *Journal of Human Resources*, 25, 253–266.

Kremer, M., and D. M. Levy (2003): "Peer Effects and Alcohol Use Among College Students," NBER Working Paper 9876.

Li, H., and T. Li (Forthcoming): "The Gender Difference of Peer Influence in Higher Education," *Economics of Education Review*.

Loury, L., and D. Garman (1995): "College Selectivity and Earnings," *Journal of Labor Ec0nomics*, 13, 289–308.

Lyle, D. 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, C. F. (1993): "Identification and Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 60(3), 531–42.

Moffitt, R. A. (2001): "Policy Interventions, Low-Level Equilibria, and Social Interactions," in *Social Dynamics*, ed. by S. Durlauf, and H. P. Young. Brookings Institution Press, Washington, DC.

Mueller, R. (1988): "The Impact of College Selectivity on Income for Men and Women," *Research in Higher Education*, 29, 175–191.

Princeton Review (2007): "Colleges and Careers," Princeton Review, Accessed on 25 August 2006 at: http://www.princetonreview.com/college/default.asp.

Sacerdote, B. L. (2001): "Peer Effects with Random Assignment: Results for Dartmouth Roommates," *Quarterly Journal of Economics*, 116(2), 681–704.

Siegfried, J. J., and M. A. Gleason (2006): "Academic Roommate Peer Effects," Working Paper.

Stinebrickner, R., and T. R. Stinebrickner (2006): "What can be learned about peer effects using college reoomates? Evidence from new survey data and students form disadvantaged backgrounds," *Journal of Public Economics*, 90(8-9), 1435–54.

Vigdor, J., and T. Nechyba (2004): "Peer Effects in North Carolina Public Schools," Unpublished Manuscript.

Weinberg, B. A. (2005): "Social Interactions and Endogenous Association," Unpublished Manuscript.

Zimmerman, D. 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
Grade Point Average (GPA) (freshman fall semester)	3407	2.88	0.62	0.28	4.00
SAT Math	3489	665.47	63.88	440.00	800.00
SAT Verbal	3489	631.95	67.00	330.00	800.00
Academic Composite	3488	1,282.41	211.99	623.00	2,067.00
Fitness Score	3489	459.70	96.88	215.00	745.00
Leadership Composite	3490	1,724.16	182.42	900.00	2,370.00
Black	3490	0.05	0.22	0	1
Hispanic	3490	0.06	0.24	0	1
Asian	3490	0.05	0.23	0	1
Female	3490	0.18	0.38	0	1
Recruited Athlete	3490	0.28	0.45	0	1
Military Preparatory School	3490	0.21	0.41	0	1
Freshman Roommate SAT Math (mean if two)	2170	665.95	55.88	460.00	800.00
Freshman Roommate SAT Verbal (mean if two)	2170	631.11	59.47	350.00	800.00
Freshman Roommate Academic Composite (mean if two)	2170	1,285.90	188.05	623.00	2,067.00
Freshman Roommate Fitness Score (mean if two)	2171	458.07	83.81	245.00	735.00
Freshman Roommate Leadership Composite (mean if two)	2171	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

Variable	Class/year	Class of 2000	Class of 2001	Class of 2002	Class of 2003	Class of 2004	Class of 2005	Class of 2006	Class of 2007
Academic Composite	Freshman Roommate	NA	NA	-0.104	NA	-0.051	-0.059	0.050	0.018
				(0.084)		(0.067)	(0.065)	(0.059)	(0.064)
	Freshman Squadron	NA	NA	-1.668***	NA	-1.029**	-0.116	0.032	-0.165
				(0.467)		(0.412)	(0.325)	(0.229)	(0.238)
	Sophomore Squadron	-0.186	-0.072	-0.020	-1.477***	-0.304	-0.117	-0.017	-0.060
		(0.313)	(0.117)	(0.250)	(0.389)	(0.226)	(0.288)	(0.166)	(0.240)
SAT Math	Freshman Roommate	NA	NA	-0.122*	NA	-0.050	-0.071	-0.017	0.080
				(0.071)		(0.063)	(0.057)	(0.074)	(0.069)
	Freshman Squadron	NA	NA	-0.420	NA	-0.237	0.255*	-0.055	-0.333
				(0.319)		(0.327)	(0.146)	(0.364)	(0.325)
	Sophomore Squadron	-0.838	-0.088	-0.154	-0.376	-0.042	0.120	-0.399	-0.532*
		(0.572)	(0.234)	(0.221)	(0.259)	(0.231)	(0.206)	(0.319)	(0.281)
SAT Verbal	Freshman Roommate	NA	NA	-0.012	NA	-0.114*	-0.104	-0.038	-0.036
				(0.052)		(0.058)	(0.064)	(0.069)	(0.073)
	Freshman Squadron	NA	NA	-0.247	NA	-1.335***	-0.418	-0.040	-0.578
				(0.294)		(0.481)	(0.266)	(0.194)	(0.355)
	Sophomore Squadron	-0.641	-0.054	0.174	-0.382	-0.490	-0.007	-0.080	-0.712
		(0.419)	(0.246)	(0.168)	(0.274)	(0.323)	(0.309)	(0.312)	(0.449)
Leadership Composite	Freshman Roommate	NA	NA	-0.037	NA	-0.012	-0.007	0.061	0.001
				(0.086)		(0.064)	(0.063)	(0.078)	(0.055)
	Freshman Squadron	NA	NA	-0.414	NA	-0.555	-0.574	0.038	0.094
				(0.296)		(0.448)	(0.383)	(0.222)	(0.224)
	Sophomore Squadron	-0.359	-0.011	-1.005**	-0.230	-0.033	0.051	-0.062	-0.124
		(0.249)	(0.189)	(0.477)	(0.214)	(0.254)	(0.193)	(0.220)	(0.270)
Fitness Score (CFT)	Freshman Roommate	NA	NA	-0.120**	NA	-0.047	0.073	-0.024	0.142**
				(0.058)		(0.062)	(0.068)	(0.054)	(0.059)
	Freshman Squadron	NA	NA	-1.192***	NA	-1.392***	-0.110	-0.0004	-0.213
				(0.438)		(0.493)	(0.248)	(0.184)	(0.267)
	Sophomore Squadron	-0.234	-0.424*	-0.239	-0.703*	-0.094	-0.002	-0.432	-0.289
		(0.293)	(0.243)	(0.242)	(0.378)	(0.222)	(0.226)	(0.386)	(0.280)

Each coefficient represents a separate regression where the individual (pre-treatment) characteristic is regressed on the peer characteristic. No other controls are included in each regression. * 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 GPA on Roommate and Squadron Pre-treatment Characteristics – reduced form estimation

Variable	1	2	3	4	5
Roommate SAT Verbal	0.009				-0.001
	(0.021)				(0.022)
Roommate SAT Math	-0.017				-0.015
	(0.023)				(0.023)
Roommate Academic Composite	0.001				0.001
	(0.005)				(0.006)
Roommate Fitness Score	0.016				0.014
	(0.014)				(0.014)
Roommate Leadership Composite	0.013**				0.013**
	(0.006)				(0.006)
Peer SAT Verbal		0.348***		0.406***	0.448***
(other freshmen in squadron)		(0.117)		(0.111)	(0.144)
Peer SAT Math		-0.106		-0.080	-0.081
(other freshmen in squadron)		(0.107)		(0.109)	(0.144)
Peer Academic Composite		-0.025		-0.030	-0.034
(other freshmen in squadron)		(0.036)		(0.034)	(0.046)
Peer Fitness Score		0.139**		0.171**	0.153*
(other freshmen in squadron)		(0.066)		(0.070)	(0.081)
Peer Leadership Composite		0.035		0.037	0.030
(other freshmen in squadron)		(0.046)		(0.037)	(0.058)
Sophomore Class SAT Verbal			0.197	0.229*	0.284**
			(0.134)	(0.119)	(0.129)
Sophomore Class SAT Math			0.082	0.124	0.176
			(0.132)	(0.133)	(0.150)
Sophomore Class Academic Composite			-0.023	-0.004	-0.008
			(0.031)	(0.029)	(0.036)
Sophomore Class Fitness Score			-0.001	-0.033	-0.060
			(0.085)	(0.076)	(0.092)
Sophmore Class Leadership Composite			-0.032	-0.012	-0.075
			(0.041)	(0.041)	(0.045)
Junior Class SAT Verbal			-0.124	-0.006	-0.013
			(0.127)	(0.115)	(0.138)
Junior Class SAT Math			-0.012	-0.002	0.112
			(0.124)	(0.122)	(0.152)
Junior Class Academic Composite			-0.003	-0.001	0.010
			(0.032)	(0.032)	(0.040)
Junior Class Fitness Score			0.122	0.085	0.097
			(0.077)	(0.077)	(0.098)
Junior Class Leadership Composite			0.056**	0.075***	0.104***
			(0.026)	(0.026)	(0.038)

Table 3: continued

Senior Class SAT Verbal			0.027	-0.033	0.051
			(0.097)	(0.106)	(0.126)
Senior Class SAT Math			0.060	0.035	-0.082
			(0.138)	(0.131)	(0.162)
Senior Class Academic Composite			-0.028	-0.046	-0.019
			(0.028)	(0.030)	(0.040)
Senior Class Fitness Score			0.011	0.012	0.067
			(0.077)	(0.082)	(0.107)
Senior Class Leadership Composite			-0.025	-0.026	-0.045
			(0.040)	(0.038)	(0.048)
SAT Verbal	0.059***	0.068***	0.065***	0.070***	0.065***
(own)	(0.020)	(0.016)	(0.016)	(0.016)	(0.020)
SAT Math	0.240***	0.260***	0.262***	0.262***	0.238***
(own)	(0.025)	(0.018)	(0.018)	(0.018)	(0.024)
Academic Composite	0.109***	0.109***	0.110***	0.109***	0.109***
(own)	(0.005)	(0.004)	(0.004)	(0.004)	(0.005)
Fitness Score	0.045***	0.050***	0.047***	0.051***	0.048***
(own)	(0.012)	(0.010)	(0.010)	(0.010)	(0.012)
Leadership Composite	0.001	0.002	0.002	0.002	0.002
(own)	(0.007)	(0.005)	(0.005)	(0.005)	(0.007)
Observations	2,166	3,404	3,404	3,404	2,166
R^2	0.3409	0.3454	0.3463	0.3507	0.3551
F-statistic (5, 107): roommate variables	1.53				1.35
F-statistic (5, 107): peer variables		2.32**		3.31***	2.46**
F-statistic (15, 107): upperclass variables			1.93**	2.12**	2.77**
F-statistic (20, 107): peer and upperclass variables				2.08***	2.38**
F-statistic (25, 107): roommate, peer, and upperclass					2.73***
Control Variables	graduation class	graduation class	graduation class	graduation class	graduation class

* 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 individua-level controls for students who are black, Hispanic, Asian, female, recruited athlete, and attended a preparatory school.

Table 4: Peer Falsification Tests

Variable	1	2	3	4
Outcome	False Peer 1		False Peer 2	
Peer SAT Verbal	0.306	0.050	0.048	0.039
	(0.222)	(0.173)	(0.100)	(0.088)
Peer SAT Math	-0.059	-0.024	-0.047	-0.020
	(0.211)	(0.164)	(0.119)	(0.086)
Peer Academic Composite	-0.086	-0.056	-0.012	-0.020
	(0.063)	(0.054)	(0.039)	(0.029)
Peer Fitness Score	0.025**	0.249**	0.025	0.019
	(0.113)	(0.104)	(0.073)	(0.059)
Peer Leadership Composite	-0.132*	0.016	0.024	0.035
	(0.077)	(0.071)	(0.036)	(0.031)
Observations	2,166	3,367	2,166	3,367
R^2	0.3443	0.3446	0.3279	0.3434
F-statistic (5, 109): peer variables	2.37**	1.28	0.18	0.33
Control Variables	roommates peer variables, graduation class	graduation class	roommates peer variables, graduation class	graduation class

* 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 peer group. All specifications include individual-level controls for SAT-v, SAT-m, academic composite, fitness score, leadership composite, black, Hispanic, Asian, female, recruited athlete, and attended a preparatory school.

Table 5: Peer Effects by Course Type

Variable	1	2	3	4	5
Outcome	Math & Science	Humanities and Social Science	Foreign Language	Physical Education	Military Studies
Peer SAT Verbal	0.672***	0.435***	-0.111	-0.005	0.289**
(other freshmen in squadron)	(0.128)	(0.126)	(0.192)	(0.086)	(0.132)
Peer SAT Math	-0.083	0.128	0.223	-0.022	-0.063
(other freshmen in squadron)	(0.134)	(0.128)	(0.150)	(0.072)	(0.120)
Peer Academic Composite	-0.083*	-0.060	-0.080	0.020	-0.042
(other freshmen in squadron)	(0.047)	(0.041)	(0.051)	(0.026)	(0.043)
Peer Fitness Score	0.177**	0.044	-0.116	0.050	0.154**
(other freshmen in squadron)	(0.074)	(0.094)	(0.122)	(0.056)	(0.078)
Peer Leadership Composite	0.020	-0.016	0.182***	0.019	0.122**
(other freshmen in squadron)	(0.055)	(0.040)	(0.064)	(0.030)	(0.049)
Observations	13,093	5,726	1,906	3,367	3,367
R^2	0.3838	0.3604	0.3692	0.5775	0.3827
Control Variables	roommates peer variables, course by section fixed effects	roommates peer variables, course by section fixed effects	roommates peer variables, course by section fixed effects	roommates peer variables, course by section fixed effects	roommates peer variables, course by section fixed effects

* 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 peer group. All specifications include individual-level controls for SAT-v, SAT-m, academic composite, fitness score, leadership composite, black, Hispanic, Asian, female, recruited athlete, and attended a preparatory school.

Table 6: Persistence in the Freshman Peer Group Effects

Variable	1	2	3	4
Outcome	Freshman Grades	Sophomore Grades	Junior Grades	Senior Grades
Peer SAT Verbal	0.519***	0.230**	0.200*	0.176**
(other freshmen in squadron)	(0.100)	(0.102)	(0.104)	(0.079)
Peer SAT Math	-0.016	0.188*	0.075	-0.017
(other freshmen in squadron)	(0.110)	(0.101)	(0.097)	(0.073)
Peer Academic Composite	-0.067*	-0.079***	-0.033	-0.037*
(other freshmen in squadron)	(0.038)	(0.027)	(0.026)	(0.021)
Peer Fitness Score	0.126**	0.035	0.013	-0.061
(other freshmen in squadron)	(0.064)	(0.067)	(0.062)	(0.062)
Peer Leadership Composite	0.022	-0.020	-0.020	0.044
(other freshmen in squadron)	(0.042)	(0.048)	(0.042)	(0.035)
Observations	27,113	26,160	25,482	22,730
R^2	0.3668	0.3939	0.4287	0.4386
Control Variables	roommates peer variables, course by section fixed effects	roommates peer variables, course by section fixed effects, current squadron by year fixed effect	roommates peer variables, course by section fixed effects, current squadron by year fixed effect	roommates peer variables, course by section fixed effects, current squadron by year fixed effect

* 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 peer group. All specifications include individual-level controls for SAT-v, SAT-m, academic composite, fitness score, leadership composite, black, Hispanic, Asian, female, recruited athlete, and attended a preparatory school.