DISCUSSION PAPER SERIES

First Generation College Students and Peer Effects

Michael S. Kofoed
Todd R. Jones

DISCUSSION PAPER SERIES

First Generation College Students and Peer Effects

Michael S. Kofoed
United States Military Academy and IZA

Todd R. Jones
Mississippi State University, CESifo and IZA

ABSTRACT

First Generation College Students and Peer Effects[*]

Higher education policymakers are concerned about the success of first-generation college students. In this study, we investigate one potential factor that may influence outcomes: first-generation students' peers. To mitigate common biases that may arise when estimating peer effects, we leverage the assignment of roommates at The United States Military Academy (West Point). We do not find evidence that being exposed to a roommate(s) with a one standard deviation higher English SAT score impacts first-semester English grades for first-generation students. Our findings for math are inconclusive, with at best suggestive evidence of a small, positive effect.

Corresponding author:
Michael S. Kofoed
Department of Social Sciences
The United States Military Academy
607 Cullum Road, West Point
New York
USA

E-mail: michael.kofoed@westpoint.edu

1 Introduction

First-generation college students fare worse than their non-first-generation counterparts on a number of outcomes. For instance, first-generation college students are 14 percentage points less likely to graduate (DeAngelo et al., 2011).[1] Because lower graduation rates are associated with lower earnings (Ost, Pan and Webber, 2018), first-generation students may also face greater difficulty paying back their student loans. Additionally, first-generation college students are more likely to under-match academically by enrolling in colleges with lower average entrance exam scores than their own (Smith, Pender and Howell, 2013; Startz, 2022). They are also more likely to come from underrepresented groups (Toutkoushian, Stollberg and Slaton, 2018). Therefore, ensuring access, persistence, and graduation of first-generation students is a major concern for higher education policymakers (Terenzini et al., 1996; Bound and Turner, 2011; Startz, 2022). As a result, many colleges have implemented policies to assist first-generation students, including first-generation-only dorms, mentoring and tutoring programs (Sullivan et al., 2021), and peer advising programs (Ellis and Gershenson, 2020).

The results from various financial and administrative interventions are mixed. Angrist, Autor and Pallais (2022) randomize financial aid awards to low income, minority, first-generation college students in Nebraska and find that receipt of aid increased four-year college going among these students. Angrist, Lang and Oreopoulos (2009) randomize institutionally sponsored academic support services and find that this intervention boosted grades and academic standing for first generation college students. Ortagus and Kramer, II (2022) demonstrate that no-loan promises for graduate school admissions also encourages more first-generation college student enrollment. In contrast, Gurantz et al. (2021) randomize low-income and first-generation students into an information nudge treatment to encourage them to enroll in selective colleges and find no effect on college choice. Turner and Gurantz (2023) randomize access to college coaching for a sample of

---

[1]We define first-generation students as those whose parents did not attend college.

students who had dropped out of college, which was heavy on first-generation students, and find no effect on re-enrollment.

In this paper, we focus on an understudied factor that may affect the academic success of first-generation students: their peers. We investigate whether first-generation students perform better when paired with higher-ability roommate(s). Given the policy concern for first-generation college students, the literature regarding the effects of peers on their academic success is limited. Garlick (2018) uses a tracking mechanism at a South African university to demonstrate that students with low parental education background are particularly responsive to a university assigning them high ability peers. However, Azzam, Bates and Fairris (2022) randomize students at a large, four-year public institution into a freshman learning community and find no effect on retention.

The effect of peers on first-generation students' academic success is theoretically ambiguous. Because their parents did not attend college, first-generation students may have less guidance on developing the study habits and soft skills necessary to succeed in their post-secondary studies (Adamecz-Völgyi, Henderson and Shure, 2020). If their peers choose to help them, then they (their peers) may have a positive effect. However, first-generation students may also face social stratification and interact less with their peers, resulting in lower transfer of social capital (Carrell, Sacerdote and West, 2013). The academic quality of a first-generation college student's peers could also influence student success.[2]

Our focus is on the effects of the assignment of peer quality (as measured by SAT scores earned before enrollment at the Academy) on outcomes for first-generation students. Our results may be helpful as some universities and high schools design first year experience programs to aid in the transition to college that may include peer mentorship, special class sections, and administrative support (Kezar and Kitchen, 2020; Azzam,

---

[2]Other studies have studied peer effects with a focus on higher-performing peers such as Sacerdote (2001), Carrell, Fullerton and West (2019), Bursztyn and Jensen (2017), Mouganie and Wang (2020), and Modena, Rettore and Tanzi (2022).

Bates and Fairris, 2022; Bettinger and Baker, 2014). However, estimating the effectiveness of these program is difficult because participation is generally non-random and students may opt into the program because they are motivated by unobservable ability or desire to succeed. Universities are also reluctant to randomly assign roommates or social networks out of a desire for fairness or reluctance to help some students but not others. All of these factors make empirically estimating the causal effect of peer effects difficult (Manski, 1993).

In our study, we focus on The United States Military Academy located in West Point, New York, where first-year students are assigned roommates; we discuss the assignment approach in detail later in the paper. This characteristic is important as it helps us address the issue of selection bias, which arises when people select their peer groups based on hidden factors associated with the outcome of interest. Furthermore, we utilize a pre-determined proxy for ability, namely the SAT score, to address the reflection problem, which arises when the variable on the right-hand side is measured after exposure to peers, and it is uncertain which peer influenced whom.

We document several findings. First, we examine roommate peer effects in general, without taking into account first-generation status, and find no evidence that students exposed to a roommate[3] with a math (English) SAT score one standard deviation higher achieve better first semester math (English) grades.[4] We do not detect a peer effect for *first generation students* who have a higher-ability English roommate. Our results for math are inconclusive, with at best suggestive evidence of a positive effect. For English, we can rule out effect sizes greater than 0.08 standard deviations. An empirical p-value analysis fails to provide evidence of an effect. Additionally, we present results by one's own SAT score and, for math, find larger coefficients for the lowest-ability cadets, but the coefficients are

---

[3]We use the singular roommate in writing, even though cadets can have more than one roommate.

[4]Lyle (2009) and Lyle (2007) also fail to detect an effect in first-year grades at West Point. He considers companies, which are a larger groups of peers. He argues that this may be due to the randomization process to company being conditional on SAT scores, limiting variation across peer groups. In our setting, we focus on roommates, which should have greater variation as they are assigned only conditional on gender, company, and class cohort.

not statistically significant. Next, we consider subsample analyses and do not find an effect for most categories. However, for math, we do find positive and significant effects for cadets who are Black and who attended a military preparatory school. Lastly, we do not find statistically significant evidence that a higher-ability peer affects first generation students in other mandatory first-year courses.

We contribute to the emerging literature on the impact of peer parental education on a student's academic performance. While previous studies such as Fruehwirth and Gagete-Miranda (2019) and Chung (2020) find that parental education matters for K-12 students in the U.S., they use non-random assignment. We argue that our use of roommate assignment improves on their identification strategy. In contrast, two other studies, Chung and Zou (2020) and Weng (2021), focus on middle school students in China and use random assignment to classrooms, while we focus on assignment to roommates among college students.

Our study distinguishes itself from these four papers by focusing on first-generation college students. Furthermore, we examine the role of peer ability, rather than parental ability as proxied by degree attainment.[5] Our research fills a gap in the literature by exploring the impact of peers on the academic achievement of college students, particularly those who are first-generation students.

2 Setting

West Point is a highly selective liberal arts college located north of New York City in the Hudson Valley. It is operated by the United States Army. Cadets at West Point are not permitted to be married or have dependents. Moreover, they do not pay any tuition or fees and receive a moderate stipend during their four years of residency. After graduation, they are commissioned as Army officers and are required to complete a service

---

[5]The role of peer ability has been the focus of previous studies such as Lyle (2007), Lyle (2009), Carrell, Fullerton and West (2019), Jain and Kapoor (2015), and many others

obligation of five years of active duty and three years in the Army Reserves.

Before the start of their freshman year, cadets at West Point undergo basic training, following which they are randomly assigned to one of thirty-six companies. A company comprises approximately 100 cadets, ranging from freshmen to seniors. In addition, West Point assigns a tactical officer to each company who is an active-duty military officer responsible for overseeing cadet development in the military, academic, and physical fitness domains (Kofoed and mcGovney, 2019).

After being assigned to a company, West Point assigns a cadet to one or more roommates from the same company, with the restriction that all roommates are of the same gender; a cadet keeps the same roommates for at least the first semester.[6] While cadets are assigned to roommates, there is no official policy on how this assignment is to be performed, though assignment within a company is under the purview of a company's tactical officer.[7] According to the Brigade Tactical Officer, who oversees all tactical officers, "[Tactical officers] generally randomly assign roommates for [first-year cadets]."[8] Thus, while we do not claim that all roommate assignment during the first semester is random, much of it appears to be. Below, we test for (as good as) random assignment and proceed to estimate the effects of high-performing roommates on the academic success of first-generation college students.

In addition to assigned roommates, West Point cadets takes a largely uniform curriculum of coursework during their first two years of study. With limited exceptions, first-year cadets take identical math and English courses, which are the focus of our

---

[6]The West Point Association of Graduates Parent Handbook 2021-22 states that cadets are randomly assigned to roommates during basic training. It then says: "Between the end of CBT [basic training] and the start of the academic year, cadets will be assigned to rooms in their permanent company areas. Cadets are generally assigned two or three to a room…Once assigned rooms, Plebes retain the assignments from August through December." See https://www.impelcreative.com/publications/wpaog/parenthandbook21-20/28/. Accessed May 12, 2023.

[7]The Cadet SOP (page 90) does state that: "TAC Teams and Cadet Company Commanders will manage roommate changes and will ensure no Fourth-, Third- or Second-Class Cadets remain together for two consecutive semesters. The RTO retains authority to approve deviations from this standard."

[8]Personal correspondence on May 12, 2023. In addition, anecdotally, we have heard that cadets are randomly-assigned to roommates for their first semester.

study. In the first semester of the first year, the typical cadet takes Mathematical Modeling/Introduction to Calculus (MA103)[9] and Introduction to Composition (EN101).[10] In addition, cadets are required to take several other courses in their first year, including psychology, chemistry, geography, and IT;[11] we consider these in an auxiliary analysis.

West Point (and the Air Force and Naval Academies) randomly assigns cadets to their instructors (when more than one teaches a course) (Carrell, Page and West, 2010), class hour (Haggag et al., 2021), and semester timing (for select courses, excluding the math and English courses that are our focus) (Patterson, Pope and Feudo, Forthcoming). Since cadets cannot choose when they take their classes or their instructors, West Point requires that first-year curriculum and graded events across instructors within the same course are identical. This policy implies that instructors teach and assign the same math and English to typical first-year roommates simultaneously.

3 Data

We obtain cadet data from the Office of Economic and Manpower Analysis, which is an organization located at West Point that advises the Army on personnel matters. We consider the 2014 and 2016 to 2023 graduating classes, where the year indicates the year of graduation. We exclude the 2015 class due to insufficient data to reconstruct roommate assignments for nearly all cadets. In practice, almost all of the 2017 class is also excluded due to missing data.

We supplement the student data with parental education data from The Freshman Survey, a survey conducted by the Higher Education Research Institution at UCLA. We

---

[9]Two notable exceptions are that some instead take honors math (MA 153), and that others take remedial math (MA 100) in the fall and MA 103 in the spring.

[10]Two notable exceptions are that some take this course in the spring and others take honors English (EN 151, honors composition).

[11]See https://courses.westpoint.edu/static/AcademicProgram.htm. Not all of these are always taken in the first semester—they may be taken in the second semester.

define a student to be first generation if neither of his/her parents attended college.[12] We test sensitivity to this definition in a robustness check.

Our main outcome variables are the semester grade that cadets receive in their MA103 or 153 math course, and in their EN101 or 151 English course. To facilitate interpretation, we standardize these scores.[13][14] In our main results, we group MA103 and MA153, as well as EN101 and 151, together. Additionally, we present results after restricting the analysis to MA103 and EN101, respectively.

To reconstruct roommate assignments, we rely on key check-out records. We assume that cadets from a given first year, defined as the year they started at West Point and were assigned to their first-year roommates, are roommates if they are coded as being in the same room facility and same room number. To determine the first year, we use the minimum of the graduation year (minus three) and the earliest year in which the cadet took MA 103/153 or EN 101/151. If a cadet is missing both math and English, we use their graduation year (minus three). This approach should produce a more accurate identification of the first year, as some cadets may take longer than four years to graduate.

We exclude cadets who are missing SAT scores or information on roommate assignments, as well as those who, after these restrictions, have zero or more than five roommates. Additionally, we remove instances where a room is coded as having both male and female cadets, which is not possible in practice but may occur due to the need to construct roommate assignments. Once we construct our sample, we compute standardized leave-out mean roommate SAT scores (separately for math and English) for each cadet and standardize the dependent variables. We then exclude cadets who lack first-generation

---

[12]In other words, a student is considered to be first generation if neither his/her mother and father attended college (or if his/her father (mother) did not attend college if we only observe the education status of the father (mother). We consider the following groups as having attended college: "Some college," "College grad," "Some grad school," "Grad degree," and "Doctorate PhD." We consider the following groups as not having attended college: "Junior high, middle school, or less," "Grammar school or less," "Some HS," and "Postsecondary."

[13]After our main sample selections, we standardize all data, regardless of year, together. When we do subsample analyses, we do not re-standardize on the subset. Results should be interpreted accordingly.

[14]It is worth noting that dependent variables standardized using this conventional method may be underestimated (Penney, 2023).

status or have missing observations for both math and English. We implement these restrictions after constructing the leave-out mean roommate SAT variables to enhance their accuracy.

3.1 Summary Statistics

Table 1 presents sample summary statistics. The first column displays data from the full sample, while the second and third columns show information on first-generation and non-first-generation students, respectively. The final column displays the difference between these two groups. First-generation students constitute 5% of the sample and have lower SAT math scores than their roommates. They are more likely to be female, Black, Hispanic, or of another race, as well as graduates of the United States Military Academy Preparatory School.[15] Additionally, they are more likely to have a prior enlistment and lower SAT scores. We also demonstrate in Appendix Figure 2 that there is variation in the leave-out mean SAT scores of roommates.

4 Empirical Strategy

4.1 Check for Randomization

We start by testing our key identifying assumption that the assignment to roommates at West Point is as good as random (conditional on gender). To do so, we regress the math or English SAT score of cadet i's roommate (or the average leave-out score in the case of multiple roommates) on a vector of observable characteristics. We anticipate that if roommate assignment is exogenous, math and English SAT scores will be unrelated to the observable characteristics of the cadet.

Table 2 shows our findings. Columns 1-3 (4-6) use leave-out roommate mean math

---

[15]The United States Military Academy Preparatory School provides remedial courses to students who may require an additional year to develop academically before re-applying to West Point.

(English) SAT score as the dependent variable. Columns 1 and 4 include the entire sample, while Columns 2 and 5 limit to men, and Columns 3 and 6 limit to women. We present the results separately by gender since roommates are always of the same gender. We also control for selection pool (year-company-gender) fixed effects and the leave-out mean of the dependent variable, calculated at the selection pool level (Guryan, Kroft and Notowidigdo, 2009).

Only the coefficients on a cadet being prior enlisted in Column 3 (math sample) and "other race" in Columns 4 and 5 (English sample) are significant. The coefficients for own SAT math, own SAT verbal, and first-generation status are never significant. Overall, we consider the table as evidence consistent with the quasi-random assignment of roommates with respect to roommate SAT scores.

4.2 Estimating Equation

Given evidence consistent with quasi-random assignment of first-generation college students with respect to their roommates' academic ability as measured by SAT scores, the econometric model is straightforward:

$$Y_{idcgt} = \beta_0 + \beta_1 FirstGen_i + \beta_2 SAT_i + \beta_3 \overline{SAT_{j(d,t)-i}} + \beta_4 FirstGen_i \times \overline{SAT_{j(d,t)-i}}$$
$$+ SelectionPool_{cgt} + \epsilon_i \tag{1}$$

Here, individual i is of gender g (male or female) and lives in dorm room d in company c in year t. The outcome variable, Y, is the course grade. The model includes an indicator for first-generation status ($FirstGen_i$), the SAT score of the individual (SAT_i), and the average SAT score of their roommates ($\overline{SAT}$) computed over the other cadets in the same room-year ($j(d,t) - i$). Our coefficient of interest is β_4, the interaction between first-generation status and the leave-out roommate SAT score. We also control for selection pool fixed effects ($SelectionPool_{cgt}$) to account for West Point's policy of assigning

roommates within companies of the same gender and cohort.[16] We use robust standard errors and cluster at the year-company-room level.[17]

4.3 Empirical P-values

In addition to the baseline regressions, we present results using empirical p-values. We show where the coefficient of interest falls in the distribution of 5,000 coefficients obtained after synthetically randomly re-shuffling roommate assignments.

We first impose initial sample restrictions and save a baseline dataset. In the first iteration, we randomly reassign cadets within a given selection pool (company-gender-year) to rooms within that same selection pool. We then apply additional sample restrictions and standardize the variables as described earlier. We estimate our regressions (one for math and one for English) and save the coefficients. We repeat this process 5,000 times to obtain a distribution of coefficients that we can use to compare the coefficient obtained with the actual data.

5 Results

After controlling for other variables, including own SAT math score, we do not find statistically significant evidence that the first-generation student main effect is associated with a higher or lower math grade. Students with a one standard deviation higher SAT math score receive a 0.46 standard deviation better grade in the course, indicating that our proxy for math ability is working as intended. Assigning a student to a roommate(s) with a one standard deviation higher SAT math score (the other main effect) does not lead to an increase in math grade.[18] The coefficient of interest, the interaction between

---

[16]We note that we are unable to compare first-generation students with non-first-generation roommates to those with first-generation roommates due to the small number of first-generation students in our sample.

[17]As we include only pre-determined variables on the right hand side, including own SAT score, we understand that exclusion bias (Caeyers and Fafchamps, 2021) is not an issue.

[18]Appendix Table A1 shows that these results are similar when not including the interaction term.

being a first-generation student and having a higher-ability math roommate, is positive but statistically insignificant.

Panel B of Table 3 presents the corresponding results for English. Own SAT English score continues to be a strong predictor of GPA, while first generation status and roommate ability do not. Estimates for the interaction term are even closer to zero, and we can rule out effects of 0.08 standard deviations.

Figure 1, Panel A, displays the empirical p-value results for math. The distribution of coefficients is symmetrical and is centered slightly to the right of 0. The actual value is indicated by the black line and falls somewhat to the right of the mean, failing to provide strong evidence of an effect. In Panel B, the distribution for English is centered somewhat more to the right of 0 (contrary to our expectation). The actual coefficient is close to 0 and to the left of the mean, providing no evidence of an effect.

We test the robustness of our results in Table 4 by altering how we define the first-year courses and their timing. Column 1 restricts the sample to those who took the course (MA 103/153 in Panel A and EN 101/151 in Panel B) during the first (fall) semester. In Column 2, we consider only those who took the non-honors course, MA 103 (Panel A) and EN 101 (Panel B). Column 3 combines both restrictions, and in Column 4, we also require the year of the course to have been taken in the first year, as previously defined. For math in Panel A, the coefficient increases in magnitude across columns, becoming marginally significant in the final column, providing some evidence that higher-ability roommates lead first-generation students to achieve a higher first-semester math grade (in MA 103). Results are stable across columns for English.

We vary our definitions of first generation status in Appendix Table A2. In Column 1, we reproduce our baseline results, in which we define a first-generation student as someone whose parents' highest level of education is "postsecondary" education or lower (where postsecondary education is post-high school education other than college). In Column 2, a first-generation student is someone whose parents' highest level of education

is high school graduation (without pursuing "postsecondary" education or college) or lower. In Column 3, a first-generation student is someone whose parents' highest level of education is some college (but not a college degree) or lower. Although the results are similar across all columns, in Panel A (math), the standard errors decrease significantly in Column 3, but still not quite enough to detect a marginally significant effect.

We also present results from a fully-interacted model in Appendix Table A3. Utilizing the fully-interacted specification yields coefficients that are identical to those obtained from regressions conducted exclusively on a sample consisting of first generation students. The math coefficient falls from the final column 0.55 in Table 3 to 0.40 in the final column of Appendix Table A3. The results for English in the final column of both tables remain virtually unchanged.

Table 5 Panel A analyzes the baseline specification and sample (Column 4 of Table 3), but divides the sample according to own SAT score as a proxy for math ability. Column 1 replicates Column 4 of Table 3 for comparison purposes. In Column 2, we limit the sample to an SAT Math score below 640; in Column 3, we limit it to a score between 640 and 680, and in Column 4, we limit it to a score above 680. The interaction term coefficient for those with the lowest SAT math scores (Column 2) is positive but not statistically significant. This coefficient decreases as the SAT math score increases. At best, this provides weak evidence that first-generation cadets with weaker math ability benefit more from having roommates with stronger math ability than cadets with stronger math ability. However, due to the large standard errors, we cannot make a conclusive statement. Panel B performs the same analysis for English, and in no case is the coefficient statistically significant.

In Table 6, we consider various demographic groups separately, including Black cadets, male cadets, female cadets, cadets with prior military service, and cadets who attended a military preparatory school. We examine the interaction effect for Panel A (math) and find evidence that first-generation Black cadets and first-generation cadets who attended

a military preparatory school benefit from being paired with a higher-ability roommate. However, we did not detect any effects for any subgroup with English (Panel B).

Table 7 examines the grade in various courses commonly taken in the first year. Panel A presents the results when using SAT math on the right-hand side, while Panel B shows the same when using SAT English. Though the math coefficients tend to be positive and of the same magnitude as the main math results, we find no significant effects for any subject.

6   Conclusion

One perplexing issue in higher education is the success of first generation college students. Given increased social mobility of attending college for first-generation college students, the returns to education are quite high. However, these students are more likely to earn lower grades and drop-out at higher rates. One way that colleges and universities hope to retrain first-generation college students is by first year experience courses and peer mentoring programs. These programs aim to assist in the transmission of human capital from high-ability peers to assist first-generation college students, but are difficult to measure in a college setting.

To examine whether peers assist first-generation college students to do achieve higher outcomes in the classroom, we leverage roommate assignments at the United States Military Academy at West Point. This setting is advantageous because West Point assigns freshman roommates and, for the most part, these students take the same classes with the same assignments and curriculum. This setting allows us the opportunity to estimate causally the effect of roommate ability on first-generation student outcomes.

Our study suggests that peers play a limited role in influencing academic success of first-generation college students. We find no effect of roommate ability on a first-generation student's grade in an introductory English class and limited evidence of a pos-

13

itive effect for math. We perform an empirical p-value analysis and do not find evidence of an effect for either subject. We do find evidence that peers have an effect on math scores for first generation Black students and first generation students who attended a military preparatory school, with more limited evidence for first generation students with lower SAT scores.

Our findings are relevant to current concerns about first-generation college students. The lack of statistically or economically significant peer effect is important in our context because it can shine light on the efficacy of programs designed to expose first-generation college students to their peers. In our setting, we fail to obtain find strong evidence that having a roommate with a higher SAT-score induces greater academic performance for first-generation students. Higher education policymakers may want to exercise caution or temper expectations about the efficacy of similar programs.

References

Adamecz-Völgyi, Anna, Morag Henderson, and Nikki Shure. 2020. "Is 'First in Family' a Good Indicator for Widening University Participation." *Economics of Education Review*, 78: 102038.

Angrist, Joshua, Daniel Lang, and Philip Oreopoulos. 2009. "Incentives and Services for College Achievement: Evidence from a Randomized Trial." *American Economic Journal: Applied Economics*, 1(1): 136–163.

Angrist, Joshua, David Autor, and Amanda Pallais. 2022. "Marginal Effects of Merit Aid for Low-Income Students." *The Quarterly Jorunal of Economics*, 137(2): 1039–1090.

Azzam, Tarek, Michael D. Bates, and David Fairris. 2022. "Do Learning Communities Increase First Year College Retention? Evidence from a Randomized Control Trial." *Economics of Education Review*, 89: 102279.

Bettinger, Eric P., and Rachel B. Baker. 2014. "The Effects of Student Coaching: An Evaluation of a Randomized Experiment in Student Advising." *Educational Evaluation and Policy Analysis*, 36(1): 3–19.

Bound, John, and Sarah Turner. 2011. "Droupouts and Diplomas: The Divergence in Collegiate Outcomes." *Handbook of the Economics of Education*, 4: 573–613.

Bursztyn, Leonardo, and Robert Jensen. 2017. "How Does Peer Pressure Affect Educational Investments?" *The Quarterly Journal of Economics*, 98: 144–168.

Caeyers, Bet, and Marcel Fafchamps. 2021. "Exclusion bias in the estimation of peer effects." National Bureau of Economic Research.

Carrell, Scott E., Bruce I. Sacerdote, and James E. West. 2013. "From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation." *Econometrica*, 81(3): 855–882.

Carrell, Scott E., Marianne E. Page, and James E. West. 2010. "Sex and Science: How Professor Gender Perpetuates the Gender Gap." *Quarterly Journal of Economics*, 125(3): 1101–1144.

Carrell, Scott E., Richard L. Fullerton, and James E. West. 2019. "Does Your Cohort Matter? Measuring Peer Effects in College Achievement." *Journal of Labor Economics*, 27(3): 439–464.

Chung, Bobby W. 2020. "Peers' Parents and Educational Attainment: The Exposure Effect." *Labour Economics*, 64.

Chung, Bobby W., and Jian Zou. 2020. "Understanding Spillover of Peer Parental Education: Randomization Evidence and Mechanisms." *HCEO Working Paper, no. 2020-045*.

DeAngelo, Linda, Ray Franke, Sylvia Hurtado, John H Pryor, and Serge Tran. 2011. *Completing college: Assessing graduation rates at four-year institutions.* Los Angeles, CA: Higher Education Research Institute, UCLA.

Ellis, Jimmy R., and Seth Gershenson. 2020. "Gender, Peer Advising, and College Success." *Labour Economics*, 62: 101775.

Fruehwirth, Jane Cooley, and Jessica Gagete-Miranda. 2019. "Your Peers' Parents: Spillovers from Parental Education." *Economics of Education Review*, 73.

Garlick, Robert. 2018. "Academic Peer Effects with Different Group Assignment Policies: Residential Tracking versus Random Assignment." *American Economic Journal: Applied Economics*, 10(3): 345–369.

Gurantz, Oded, Jessica Howell, Michael Hurwitz, Cassandra Larson, Matea Pender, and Brooke White. 2021. "A National-Level Information Experiment to Promote Enrollment in Selective Colleges." *Journal of Policy Analysis and Management*, 40(2): 453–479.

Guryan, Jonathan, Kory Kroft, and Matthew J Notowidigdo. 2009. "Peer effects in the workplace: Evidence from random groupings in professional golf tournaments." *American Economic Journal: Applied Economics*, 1(4): 34–68.

Haggag, Kareem, Richard W. Patterson, Nolgan G. Pope, and Aaron Feudo. 2021. "Attribution Bias in Major Decisions: Evidence from the United States Military Academy." *Journal of Public Economics*, 200.

Jain, Tarun, and Mudit Kapoor. 2015. "The Impact of Study Groups and Roommates on Academic Performance." *The Review of Economics and Statistics*, 97(1): 44–54.

Kezar, Adrianna, and Joseph A. Kitchen. 2020. "Supporting First-Generation, Low-Income, and Underrepresented Students' Transitions to College Through Comprehensive and Integrated Programs." *American Behavioral Science*, 64(3): 223–229.

Kofoed, Michael S., and Elizabeth mcGovney. 2019. "The Effect of Same-Gender and Same-Race Role Models on Occupation Choice: Evidence from Randomly Assigned Roommates." *Journal of Human Resources*, 54(2): 430–467.

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.

Lyle, David S. 2009. "The Effects of Peer Group Heterogeneity on the Production of Human Capital at West Point." *American Economic Journal: Applied Economics*, 1(4): 69–84.

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

Modena, Francesca, Enrico Rettore, and Giulia Martina Tanzi. 2022. "Asymmetries in the Gender Effect of High-Performing Peers: Evidence from Tertiary Education." *Labour Economics*, 78: 102225.

Mouganie, Pierre, and Yaojing Wang. 2020. "High-Performing Peers and Female STEM Choices in School." *Journal of Labor Economics*, 38(3): 805–841.

Ortagus, Justin C., and Dennis A. Kramer, II. 2022. "The Impact of No-Loan Program Participation on the Likelihood of Graduate Student Enrolment Among Low-Income, First-Generation Students." *Education Finance and Policy*, 17(1): 81–104.

Ost, Ben, Weixiang Pan, and Douglas Webber. 2018. "The Returns to College Persistence for Marginal Students: Regression Discontinuity Evidence from University Dismissal Policies." *Journal of Labor Economics*, 36(3): 779–805.

Patterson, Richard W., Nolan G. Pope, and Aaron Feudo. Forthcoming. "Timing Matters: Evidence from College Major Decisions." *Journal of Human Resources*.

Penney, Jeffrey. 2023. "Cautions When Normalizing the Dependent Variable in a Regression as a Z-Score." *Economic Inquiry*, 61(2): 402–412.

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

Smith, Johnathan, Matea Pender, and Jessica Howell. 2013. "The Full Extent of Student-College Academic Undermatch." *Economics of Education Review*, 32: 247–261.

Startz, Richard. 2022. "First-Generation College Students Face Unique Challenges." *Brookings Institution, Brown Center Chalkboard*.

Sullivan, Zach, Ben Castleman, Gabrielle Lohner, and Eric Bettinger. 2021. "College Advising at a National Scale: Experimental Evidence from the CollegePoint Initiative." *EdWorkingPaper*, 19-213: 1–55.

Terenzini, Patrick T., Leonard Springer, Patricia M. Yaeger, Ernest T. Pascarella, and Amaury Nora. 1996. "First-Generation College Students: Characteristics, Experiences, and Cognitive Development." *Research in Higher Education*, 37: 1–22.

Toutkoushian, Robert K., Robert A. Stollberg, and Kelly A. Slaton. 2018. "Talking 'Bout My Generation: Defining 'First-Generation College Students' in Higher Education Research." *Teachers College Record*, 120(4): 1–46.

Turner, Lesley J., and Oded Gurantz. 2023. "Experimental Evidence of College Coaching on Postsecondary Re-enrollment." *EdWorkingPaper*, 23-746: 1–50.

Weng, Tianheng. 2021. "Classroom Composition and Student Academic Achievement: The Impact of Peers' Parental Education." *The B.E. Journal of Economic Analysis & Policy*, 21(1): 273–305.

7 Tables

Table 1: Summary Statistics

	(1) Full Sample mean/sd	(2) First Gen mean/sd	(3) Non-First Gen mean/sd	(4) Difference b/se
Female	0.210	0.238	0.209	-0.030
	(0.408)	(0.427)	(0.406)	[0.019]
Black	0.128	0.180	0.125	-0.055***
	(0.334)	(0.385)	(0.331)	[0.016]
Hispanic	0.091	0.178	0.086	-0.092***
	(0.287)	(0.383)	(0.280)	[0.013]
Asian	0.080	0.115	0.078	-0.037***
	(0.272)	(0.319)	(0.268)	[0.013]
Other Race	0.037	0.056	0.036	-0.020**
	(0.190)	(0.231)	(0.187)	[0.009]
Prep School Grad	0.151	0.285	0.144	-0.141***
	(0.358)	(0.452)	(0.351)	[0.017]
NCAA Athlete	0.206	0.190	0.207	0.017
	(0.405)	(0.393)	(0.405)	[0.019]
Prior Enlisted	0.038	0.107	0.034	-0.072***
	(0.192)	(0.309)	(0.182)	[0.009]
First Gen	0.054	1.000	0.000	-1.000
	(0.226)	(0.000)	(0.000)	[0.000]
Own SAT Math	648.147	615.921	649.980	34.059***
	(73.452)	(73.318)	(73.038)	[3.435]
Own SAT Verbal	630.877	591.946	633.091	41.145***
	(79.838)	(79.904)	(79.267)	[3.729]
Roommate Mean SAT Math	646.954	644.367	647.101	2.734
	(57.270)	(58.412)	(57.204)	[2.693]
Roommate Mean SAT Verbal	629.772	629.500	629.788	0.288
	(62.404)	(64.315)	(62.297)	[2.934]
Math Grade	3.054	2.870	3.064	0.194***
	(0.894)	(0.898)	(0.892)	[0.043]
English Grade	2.746	2.580	2.756	0.175***
	(0.720)	(0.753)	(0.717)	[0.034]
Observations	8883	478	8405	8883

Notes: The first column shows the full sample, the second column limits to first-generation students, and the third column limits to non-first-generation students. For each of these three columns, the mean value is shown along with the standard deviation in parentheses. The final column shows the difference between the second and third column, with the standard error in brackets. * 0.10; ** 0.05; *** 0.01.

Table 2: Balance Table for Roommate Ability

	Roommate Math SAT			Roommate English SAT		
	(1)	(2)	(3)	(4)	(5)	(6)
	Overall	Male	Female	Overall	Male	Female
Female	-0.120	0.000	0.000	0.062	0.000	0.000
	(0.157)	(.)	(.)	(0.304)	(.)	(.)
Black	-0.040	-0.037	-0.047	-0.025	-0.008	-0.089
	(0.034)	(0.040)	(0.067)	(0.035)	(0.039)	(0.072)
Hispanic	0.005	-0.013	0.079	-0.018	-0.031	0.029
	(0.036)	(0.041)	(0.074)	(0.036)	(0.041)	(0.077)
Asian	-0.007	-0.022	0.040	-0.005	-0.039	0.115
	(0.041)	(0.048)	(0.080)	(0.039)	(0.044)	(0.080)
Other Race	-0.053	-0.083	0.070	-0.100*	-0.106*	-0.067
	(0.059)	(0.068)	(0.115)	(0.054)	(0.062)	(0.113)
Prep School Grad	0.006	0.005	-0.008	-0.022	-0.022	-0.037
	(0.034)	(0.037)	(0.076)	(0.034)	(0.037)	(0.081)
NCAA Athlete	0.012	0.024	-0.025	0.009	0.021	-0.038
	(0.027)	(0.032)	(0.053)	(0.027)	(0.032)	(0.055)
Prior Enlisted	-0.051	-0.019	-0.347**	0.062	0.087	-0.137
	(0.055)	(0.059)	(0.149)	(0.058)	(0.062)	(0.171)
First Gen	0.024	0.034	-0.014	0.023	0.033	0.001
	(0.046)	(0.053)	(0.095)	(0.047)	(0.054)	(0.100)
Own SAT Math	0.010	0.003	-0.024			
	(0.023)	(0.101)	(0.095)			
Own SAT Verbal				-0.005	-0.077	-0.121
				(0.022)	(0.095)	(0.084)
Observations	8883	7015	1868	8883	7015	1868
R^2	0.120	0.068	0.258	0.124	0.074	0.288

Notes: This table shows a regression of mean roommate SAT score (math SAT in columns 1-3 and English SAT for columns 4-6) on a set of variables. Each column is a separate regression. Columns 1 and 4 are for the entire sample; columns 2 and 5 are limited to males; and columns 3 and 6 are limited to females. Each column includes year fixed effects and selection pool (year-company-gender) fixed effects. * 0.10; ** 0.05; *** 0.01. Standard errors are reported in parentheses and are clustered at the year-company-room level.

Table 3: Math and English

Panel A: Math	(1)	(2)	(3)	(4)
First Gen	0.014	-0.002	-0.005	0.026
	(0.043)	(0.042)	(0.044)	(0.043)
Own SAT Math	0.518***	0.524***	0.528***	0.458***
	(0.009)	(0.009)	(0.010)	(0.012)
Roommate Mean SAT Math	-0.004	0.002	0.009	0.008
	(0.010)	(0.010)	(0.010)	(0.010)
First Gen × Roommate Mean SAT Math	0.058	0.058	0.052	0.055
	(0.042)	(0.040)	(0.042)	(0.041)
Year FEs	No	Yes	No	No
Selection Pool FEss	No	No	Yes	Yes
Controls	No	No	No	Yes
Observations	8771	8771	8771	8771
R^2	0.265	0.290	0.338	0.361

Panel B: English	(1)	(2)	(3)	(4)
First Gen	-0.053	-0.039	-0.052	-0.038
	(0.047)	(0.046)	(0.048)	(0.047)
Own SAT Verbal	0.389***	0.385***	0.390***	0.299***
	(0.011)	(0.011)	(0.011)	(0.013)
Roommate Mean SAT Verbal	0.000	0.005	0.013	0.011
	(0.011)	(0.010)	(0.011)	(0.011)
First Gen × Roommate Mean SAT Verbal	0.023	0.017	0.005	-0.005
	(0.042)	(0.041)	(0.043)	(0.043)
Year FEs	No	Yes	No	No
Selection Pool FEs	No	No	Yes	Yes
Controls	No	No	No	Yes
Observations	8452	8452	8452	8452
R^2	0.146	0.168	0.236	0.267

Panel A (B) of this table reports the results from a regression of standardized math (English) course grade on an indicator for being a first-generation student; one's own SAT math (English) score; the roommate(s)' mean SAT math (English) score, and an interaction of first-generation status and the roommate(s)' mean SAT math (English) score. Year fixed effects, selection pool (year-company-gender) fixed effects, and controls are included as indicated at the bottom of the table. * 0.10; ** 0.05; *** 0.01. Standard errors are reported in parentheses and are clustered at the year-company-roomlevel.

Table 4: Varying Course Definitions and Timing

Panel A: Math	(1)	(2)	(3)	(4)
	Math	Math	Math	Math
First Gen	0.026	0.057	0.053	0.055
	(0.044)	(0.046)	(0.047)	(0.047)
Own SAT Math	0.465***	0.419***	0.430***	0.429***
	(0.012)	(0.015)	(0.016)	(0.016)
Roommate Mean SAT Math	0.007	0.006	0.005	0.005
	(0.010)	(0.012)	(0.012)	(0.012)
First Gen × Roommate Mean SAT Math	0.055	0.065	0.069	0.073*
	(0.041)	(0.044)	(0.043)	(0.044)
Observations	8529	6700	6458	6453
R^2	0.363	0.319	0.323	0.321
Semester	First	Any	First	First
Course	Any	MA103	MA103	MA103
Year	Any	Any	Any	First

Panel B: English	(1)	(2)	(3)	(4)
	English	English	English	English
First Gen	-0.030	-0.038	-0.030	-0.037
	(0.046)	(0.048)	(0.047)	(0.046)
Own SAT Verbal	0.301***	0.293***	0.296***	0.301***
	(0.013)	(0.013)	(0.013)	(0.013)
Roommate Mean SAT Verbal	0.010	0.010	0.008	0.008
	(0.011)	(0.011)	(0.011)	(0.011)
First Gen × Roommate Mean SAT Verbal	-0.013	-0.003	-0.010	-0.016
	(0.043)	(0.043)	(0.043)	(0.043)
Observations	8270	8298	8116	8083
R^2	0.270	0.258	0.261	0.261
Semester	First	Any	First	First
Course	Any	EN101	EN101	EN101
Year	Any	Any	Any	First

Notes: Panel A (B) of this table reports the results from a regression of standardized math (English) course grade on an indicator for being a first-generation student; one's own SAT math (English) score; the roommate(s)' mean SAT math (English) score, and an interaction of first-generation status and the roommate(s)' mean SAT math (English) score. Column 1 restricts to taking the course in the first semester. Column 2 restricts to taking MA 103 (Panel A) and EN 101 (Panel B). Column 3 imposes both restrictions. Column 4 also restricts to taking the course in the first year. Selection pool (year-company-gender) fixed effects, and controls are included. * 0.10; ** 0.05; *** 0.01. Standard errors are reported in parentheses and are clustered at the year-company-room level.

Table 5: Math and English, by Own SAT

Panel A: Math	(1)	(2)	(3)	(4)
	Overall	SAT Math<640	$640 \leqslant SATMath \leqslant 680$	SAT Math>680
First Gen	0.026	0.112*	0.009	-0.221**
	(0.043)	(0.061)	(0.096)	(0.111)
Own SAT Math	0.458***	0.452***	0.547***	0.363***
	(0.012)	(0.035)	(0.096)	(0.036)
Roommate Mean SAT Math	0.008	0.009	0.008	0.008
	(0.010)	(0.019)	(0.019)	(0.017)
First Gen × Roommate Mean SAT Math	0.055	0.084	0.078	0.027
	(0.041)	(0.061)	(0.077)	(0.118)
Observations	8771	3660	2479	2632
R^2	0.361	0.305	0.291	0.268

Panel B: English	(1)	(2)	(3)	(4)
	Overall	SAT English<640	$640 \leqslant English \leqslant 680$	SAT English>680
First Gen	-0.038	-0.068	0.104	-0.074
	(0.047)	(0.067)	(0.110)	(0.116)
Own SAT Verbal	0.299***	0.279***	0.273***	0.258***
	(0.013)	(0.025)	(0.025)	(0.027)
Roommate Mean SAT Verbal	0.011	0.003	0.033	0.012
	(0.011)	(0.018)	(0.022)	(0.020)
First Gen × Roommate Mean SAT Verbal	-0.005	0.017	0.008	-0.101
	(0.043)	(0.064)	(0.084)	(0.147)
Observations	8452	3661	2379	2412
R^2	0.267	0.291	0.347	0.288

Notes: Panel A (B) of this table reports the results from a regression of standardized math (English) course grade on an indicator for being a first-generation student; one's own SAT math (English) score; the roommate(s)' mean SAT math (English) score, and an interaction of first-generation status and the roommate(s)' mean SAT math (English) score. The first column is for the full sample. The second, third, and fourth column limit to individuals (not roommates) with a SAT math (English) score below 640, between 640 and 680, and above 680, respectively. Selection pool (year-company-gender) fixed effects, and controls are included in all columns. * 0.10; ** 0.05; *** 0.01. Standard errors are reported in parentheses and are clustered at the year-company-room level.

Table 6: Math and English, by Own Characteristics

Panel A: Math

	(1) Overall	(2) Black	(3) Male	(4) Female	(5) NCAA Athlete	(6) Prep School	(7) Prior Service
First Gen	0.026	0.070	0.072	-0.175*	0.086	0.158	0.441
	(0.043)	(0.161)	(0.048)	(0.099)	(0.113)	(0.103)	(0.311)
Own SAT Math	0.458***	0.454***	0.454***	0.471***	0.461***	0.441***	0.238*
	(0.012)	(0.051)	(0.013)	(0.027)	(0.034)	(0.040)	(0.121)
Roommate Mean SAT Math	0.008	-0.072*	0.011	-0.006	-0.014	0.000	0.169
	(0.010)	(0.043)	(0.011)	(0.022)	(0.028)	(0.033)	(0.131)
First Gen × Roommate Mean SAT Math	0.055	0.338**	0.063	-0.029	-0.025	0.280***	-0.038
	(0.041)	(0.131)	(0.047)	(0.084)	(0.107)	(0.097)	(0.354)
Observations	8771	1122	6919	1852	1799	1331	314
R^2	0.361	0.529	0.348	0.419	0.489	0.501	0.772

Panel B: English

	(1) Overall	(2) Black	(3) Male	(4) Female	(5) NCAA Athlete	(6) Prep School	(7) Prior Service
First Gen	-0.038	0.040	0.008	-0.199*	-0.208	-0.015	0.512*
	(0.047)	(0.182)	(0.053)	(0.102)	(0.130)	(0.128)	(0.297)
Own SAT Verbal	0.299***	0.309***	0.305***	0.276***	0.387***	0.314***	0.190
	(0.013)	(0.064)	(0.014)	(0.029)	(0.038)	(0.052)	(0.143)
Roommate Mean SAT Verbal	0.011	0.017	0.011	0.012	0.014	0.057	0.037
	(0.011)	(0.051)	(0.012)	(0.022)	(0.029)	(0.041)	(0.122)
First Gen × Roommate Mean SAT Verbal	-0.005	0.125	0.012	-0.048	-0.171	-0.116	-0.427
	(0.043)	(0.189)	(0.048)	(0.093)	(0.122)	(0.124)	(0.337)
Observations	8452	1115	6712	1740	1806	1324	328
R^2	0.267	0.498	0.238	0.361	0.454	0.462	0.697

Notes: Panel A (B) of this table reports the results from a regression of standardized math (English) course grade on an indicator for being a first-generation student; one's own SAT math (English) score; the roommate(s)' mean SAT math (English) score, and an interaction of first-generation status and the roommate(s)' mean SAT math (English) score. The full sample is included in column 1. The remaining columns restrict to Black students, female students, NCAA athletes, those who attended a military prep school, and those with prior military service, respectively. Selection pool (year-company-gender) fixed effects, and controls are included in all columns. * 0.10; ** 0.05; *** 0.01. Standard errors are reported in parentheses and are clustered at the year-company-room level.

Table 7: Other Required Core Courses

Panel A: Math Right Hand Side	(1)	(2)	(3)	(4)	(5)	(6)
	Math	English	Pyschology	Chemistry	Geography	IT
First Gen	0.026	-0.079	-0.065	-0.002	-0.041	0.028
	(0.043)	(0.049)	(0.046)	(0.041)	(0.047)	(0.045)
Own SAT Math	0.458***	0.198***	0.341***	0.496***	0.354***	0.439***
	(0.012)	(0.013)	(0.012)	(0.011)	(0.012)	(0.011)
Roommate Mean SAT Math	0.008	0.018	0.020**	0.006	0.005	0.014
	(0.010)	(0.011)	(0.010)	(0.009)	(0.010)	(0.009)
First Gen × Roommate Mean SAT Math	0.055	0.022	0.059	0.057	0.044	0.031
	(0.041)	(0.045)	(0.044)	(0.042)	(0.043)	(0.043)
Observations	8771	8452	8797	8356	8542	8436
R^2	0.361	0.236	0.304	0.402	0.314	0.379

Panel B: English Right Hand Side	(1)	(2)	(3)	(4)	(5)	(6)
	English	Math	Pyschology	Chemistry	Geography	IT
First Gen	-0.038	-0.017	-0.036	-0.049	-0.041	0.010
	(0.047)	(0.046)	(0.045)	(0.045)	(0.048)	(0.047)
Own SAT Verbal	0.299***	0.194***	0.358***	0.246***	0.276***	0.296***
	(0.013)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Roommate Mean SAT Verbal	0.011	0.005	0.009	0.005	0.003	-0.001
	(0.011)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
First Gen × Roommate Mean SAT Verbal	-0.005	-0.001	0.018	0.047	-0.006	0.009
	(0.043)	(0.047)	(0.045)	(0.046)	(0.043)	(0.046)
Observations	8452	8771	8797	8356	8542	8436
R^2	0.267	0.253	0.314	0.287	0.283	0.314

Notes: Panel A (B) of this table reports the results from a regression of course grades in different subjects on an indicator for being a first-generation student; one's own SAT math (English) score; the roommate(s)' mean SAT math (English) score, and an interaction of first-generation status and the roommate(s)' mean SAT math (English) score. In Panel A, the courses are math in column 1, English in column 2, psychology in column 3, chemistry in column 4, geography in column 5, and information technology in column 6. The columns are the same for Panel B, but the first two columns are swapped. Selection pool (year-company-gender) fixed effects, and controls are included in all columns. * 0.10; ** 0.05; *** 0.01. Standard errors are reported in parentheses and are clustered at the year-company-room level.

8 Figures

Figure 1: Empirical p-values

(a) Math (b) English
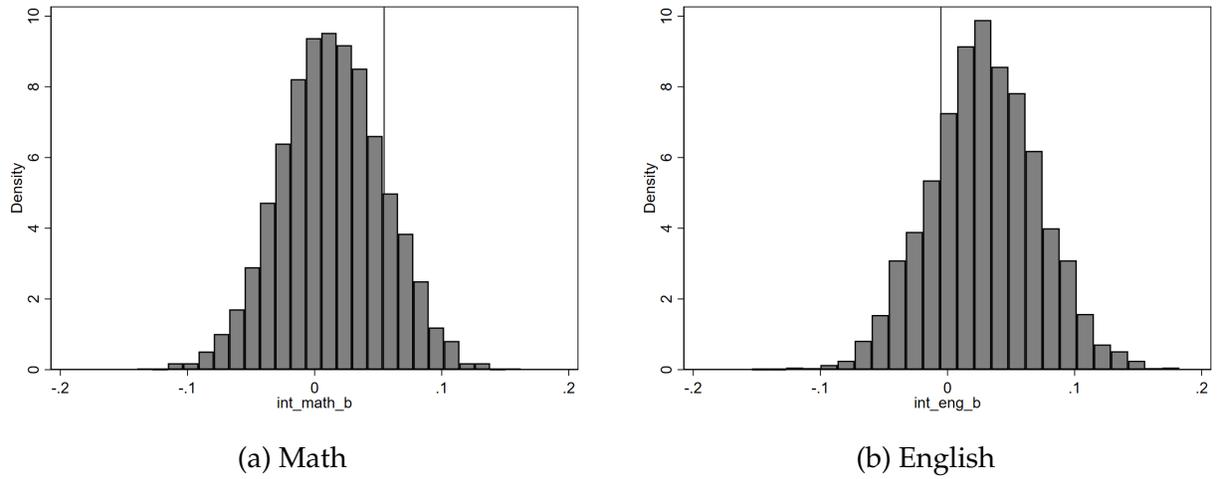
Notes: This figure shows empirical p-values for math in Panel A and English in Panel B. The histograms plot the coefficients (on the interaction term) obtained after randomizing cadets within their actual company to synthetic roommates. The actual coefficient is indicated by the black line.

A   Appendix Tables

Table A1: Math and English, No Interaction

Panel A: Math	(1)	(2)	(3)	(4)
First Gen	0.012	-0.004	-0.006	0.025
	(0.043)	(0.042)	(0.044)	(0.043)
Own SAT Math	0.518***	0.524***	0.528***	0.458***
	(0.009)	(0.009)	(0.010)	(0.012)
Roommate Mean SAT Math	-0.001	0.005	0.011	0.010
	(0.010)	(0.010)	(0.010)	(0.010)
Year FEs	No	Yes	No	No
Selection Pool FEs	No	No	Yes	Yes
Controls	No	No	No	Yes
Observations	8771	8771	8771	8771
R^2	0.265	0.290	0.338	0.361

Panel B: English	(1)	(2)	(3)	(4)
First Gen	-0.053	-0.039	-0.052	-0.038
	(0.047)	(0.046)	(0.048)	(0.047)
Own SAT Verbal	0.389***	0.385***	0.390***	0.299***
	(0.011)	(0.011)	(0.011)	(0.013)
Roommate Mean SAT Verbal	0.002	0.006	0.014	0.011
	(0.010)	(0.010)	(0.011)	(0.011)
Year FEs	No	Yes	No	No
Selection Pool FEs	No	No	Yes	Yes
Controls	No	No	No	Yes
Observations	8452	8452	8452	8452
R^2	0.146	0.168	0.236	0.267

Notes: Panel A (B) of this table reports the results from a regression of standardized math (English) course grade on an indicator for being a first-generation student; one's own SAT math (English) score; and the roommate(s)' mean SAT math (English) score. Year fixed effects, selection pool (year-company-gender) fixed effects, and controls are included as indicated at the bottom of the table. Notes: * 0.10; ** 0.05; *** 0.01. Standard errors are reported in parentheses and are clustered at the year-company-room level.

Table A2: Math and English, Varying Definitions of First Generation

Panel A: Math	(1)	(2)	(3)
First Gen	0.026	0.010	-0.047
	(0.043)	(0.049)	(0.031)
Own SAT Math	0.458***	0.457***	0.456***
	(0.012)	(0.012)	(0.012)
Roommate Mean SAT Math	0.008	0.008	0.005
	(0.010)	(0.010)	(0.010)
First Gen × Roommate Mean SAT Math	0.055	0.050	0.046
	(0.041)	(0.046)	(0.029)
Observations	8771	8771	8771
R^2	0.361	0.361	0.361
First Gen - At Most:	Postsecondary	HS Grad	Some College

Panel B: English	(1)	(2)	(3)
First Gen	-0.058*	-0.080	-0.058*
	(0.032)	(0.053)	(0.032)
Own SAT Verbal	0.298***	0.298***	0.298***
	(0.013)	(0.013)	(0.013)
Roommate Mean SAT Verbal	0.011	0.012	0.015
	(0.011)	(0.011)	(0.011)
First Gen × Roommate Mean SAT Verbal	-0.005	-0.019	-0.030
	(0.043)	(0.050)	(0.029)
Observations	8452	8452	8452
R^2	0.268	0.267	0.268
First Gen - At Most:	Postsecondary	HS Grad	Some College

Notes: Panel A (B) of this table reports the results from a regression of standardized math (English) course grade on an indicator for being a first-generation student; one's own SAT math (English) score; the roommate(s)' mean SAT math (English) score, and an interaction of first-generation status and the roommate(s)' mean SAT math (English) score. The first column defines first generation as having attended at most postsecondary education; the second column is at most high school graduate; and the third column is at most some college. Selection pool (year-company-gender) fixed effects, and controls are included in all columns. * 0.10; ** 0.05; *** 0.01. Standard errors are reported in parentheses and are clustered at the year-company-room level.

Table A3: Math and English; Fully-interacted Model

Panel A: Math	(1)	(2)	(3)	(4)
First Gen	-0.033	-0.406**	-2.197***	-0.689***
	(0.046)	(0.178)	(0.168)	(0.220)
Own SAT Math × First Gen	0.442***	0.452***	0.550***	0.483***
	(0.047)	(0.044)	(0.048)	(0.060)
First Gen × Roommate Mean SAT Math	0.058	0.067*	0.044	0.040
	(0.041)	(0.038)	(0.049)	(0.050)
Year FEs	No	Yes	No	No
Selection Pool FEss	No	No	Yes	Yes
Controls	No	No	No	Yes
Observations	8771	8771	8771	8771
R^2	0.013	0.018	0.044	0.045

Panel B: English	(1)	(2)	(3)	(4)
First Gen	-0.040	0.284*	-3.021***	-3.338***
	(0.047)	(0.161)	(0.287)	(0.350)
Own SAT Verbal × First Gen	0.392***	0.401***	0.275***	0.173*
	(0.053)	(0.054)	(0.083)	(0.090)
First Gen × Roommate Mean SAT Verbal	0.023	0.020	0.011	-0.000
	(0.041)	(0.039)	(0.047)	(0.047)
Year FEs	No	Yes	No	No
Selection Pool FEs	No	No	Yes	Yes
Controls	No	No	No	Yes
Observations	8452	8452	8452	8452
R^2	0.011	0.013	0.046	0.048

Panel A (B) of this table reports the results from a fully interacted model, in particular regression of standardized math (English) course grade on an indicator for being a first-generation student; one's own SAT math (English) score interacted with first generation status; and an interaction of first-generation status and the roommate(s)' mean SAT math (English) score. Year fixed effects, selection pool (year-company-gender) fixed effects, and controls are included as indicated at the bottom of the table; when included, the controls are also interacted with first generation status. * 0.10; ** 0.05; *** 0.01. Standard errors are reported in parentheses and are clustered at the selection pool level.

B Appendix Figures

Figure 2: Variation in Average Roommate SAT Scores
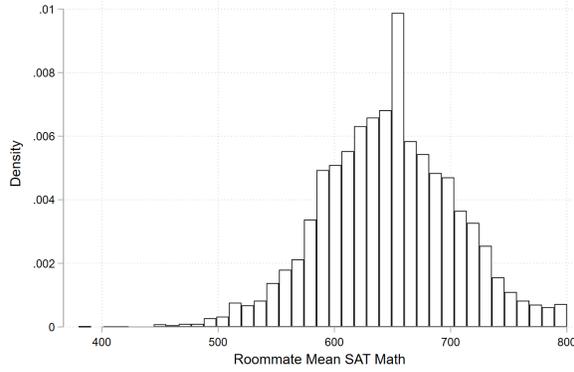
(a) Math SAT
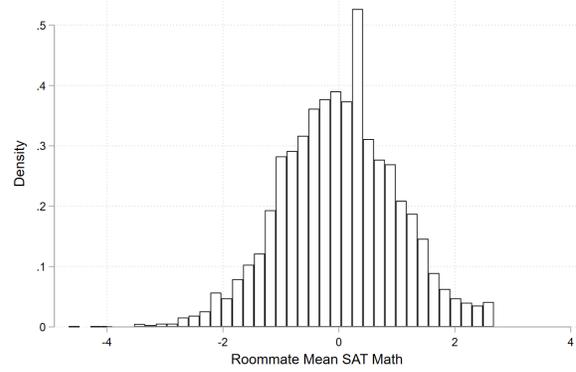
(b) Math SAT, Standardized
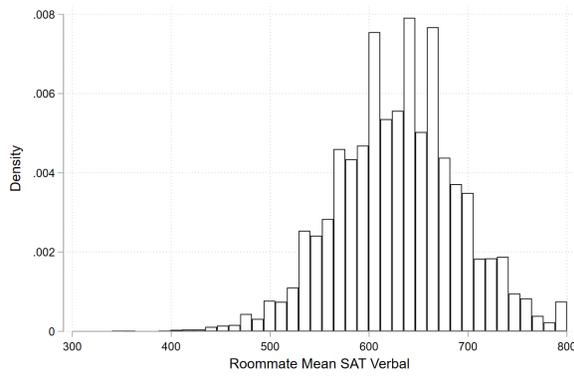
(c) English SAT
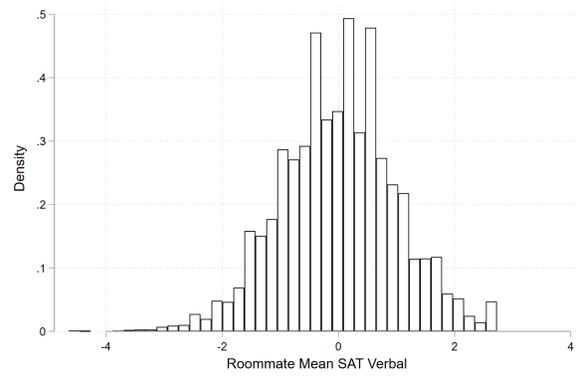
(d) English SAT, Standardized

Notes: This figure shows the variation in mean roommate SAT scores. Each observation is a cadet and the value is the mean of

that cadet's roommate(s)'s SAT score. Math is shown in Panels A and B, while English is shown in Panels C and D. The raw score

is shown in Panels A and C, and the standardized score is shown in Panels B and D.