

# The Effects of Peer Diversity on Entrepreneurship: Evidence from University Residence Halls\*

Clint Harris<sup>†</sup>

Jon Eckhardt<sup>‡</sup>

Brent Goldfarb<sup>§</sup>

September 8, 2022

## Abstract

We estimate the effects of peer academic diversity and peer gender diversity within college dorms on student entrepreneurship at a large public research university over a period of four years. To obtain causal estimates, we instrument for realized peer traits with simulated peer traits determined by the peers students would have been assigned had they been assigned to rooms by a room allocation mechanism that we select, while controlling for the expected peer traits implied by the same mechanism. We find that exposure to peer gender diversity differentially affects men and women and contributes to the gender gap in entrepreneurial intentions. We also find that exposure to peer verbal skill diversity reduces entrepreneurial ideation. Our findings suggest that residentially clustering students by gender and academic strengths can increase student entrepreneurship and reduce the entrepreneurship gender gap.

JEL Codes: M13, C52, D47, I20

Keywords: Entrepreneurship, Diversity, Peer Effects, College Majors, Simulated Instruments, Mechanism Design

---

\*We thank David Swiderski for assistance with the institutional details regarding room assignments. We thank Peiran Cheng for excellent research assistance.

<sup>†</sup>University of Wisconsin-Madison, clint.harris@wisc.edu

<sup>‡</sup>University of Wisconsin-Madison, jon.eckhardt@wisc.edu

<sup>§</sup>University of Maryland, brentg@umd.edu

# 1. Introduction

Universities play a well-known role in innovation (Hausman, 2012; Roessner, Bond, Okubo, and Planting, 2013) and are often directly involved in fostering entrepreneurship in communities (Rothaermel, Agung, and Jiang, 2007; Wright, Siegel, and Mustar, 2017; Wright and Mustar, 2019). However, despite numerous anecdotal accounts of companies formed in college dorm rooms that grow to become major employers, the scientific literature on how and when students become motivated to entrepreneurial action is scant. We focus on the role of peers in driving entrepreneurial behavior by examining how demographics of residence halls at a post-secondary institution might influence entrepreneurial intent, ideation and new firm formation. Our study is the first to ask how peer diversity in residence life may affect entrepreneurial intent. We find that proximate residential exposure to peers with different verbal skills (as measured by distance between ACT scores) reduced entrepreneurial ideation and that exposure to gender diversity contributed to gender gaps in entrepreneurial intentions. Effects of university residence hall peer exposure on entrepreneurship are particularly policy-relevant because costs of changing room assignment policies are negligible, implying that evidence-based policies informed by any statistically significant effects are essentially guaranteed to pass a cost-benefit test.

We investigate effects of peer assignments on entrepreneurial ideation and intention using data on 6,919 students who lived in dorms in their first year at the University of Wisconsin-Madison from 2016-2019, and who responded to a yearly survey of students' entrepreneurial interest and aspirations at least once. Peer diversity for a given trait is measured as the average absolute value distance between an individual's trait and the same trait for other individuals at a particular geographic distance from them. We focus in particular on differences between individuals and their roommates for traits of interest, with the notable exception of gender for which we focus on differences between individuals and their next-door neighbors. Because our identification strategy makes use of between-dorm variation in room assignments, we are also able to estimate effects of building-level characteristics.

For peer effects, we restrict our attention to diversity in student traits because such effects are particularly policy relevant. Increasing or reducing diversity exposure for some students mechanically generates the *same* effect on other students. This is at odds with the zero-sum nature of assignments based on raw peer traits. For example, increasing a student's exposure to STEM peers mechanically *decreases* other students' exposure to STEM peers, while increasing

a student’s exposure to STEM diversity (peers whose STEM statuses differ from their own) mechanically *increases* other students’ exposure to STEM diversity. It follows that main effects of peer diversity are particularly likely to inform uniformly beneficial policies, rather than ones that improve outcomes for some students at a cost to others. We also investigate the effects of dorm-level gender integration and dorm housing of entrepreneurship learning communities.

Our paper makes several contributions. First, our research is the first to examine the potential impact of room assignment policies in residence halls on student entrepreneurship, including ideation and students’ long-term interests in entrepreneurship. We find that some dimensions of peer diversity, such as neighbor gender and roommate verbal skills, negatively affect entrepreneurship. Further, we do not detect statistically significant effects of peer STEM major or peer math skill diversity on entrepreneurship. Our results are consistent with recent research that questions prior work suggesting that diversity increases entrepreneurship, such as Junkunc and Eckhardt (2009) and Khoshimov, Eckhardt, and Goldfarb (2019).

Second, we contribute evidence in favor of Learning Communities, policies employed by many universities involving residential clustering of students who share similar academic interests or other common features. Learning Communities are recommended as high-impact policies by the Association of American Colleges and Universities, they have been found effective for student academic engagement and outcomes, and they are commonly implemented by universities internationally (Zhao and Kuh, 2004; Brouwer, Flache, Jansen, Hofman, and Steglich, 2018). University of Wisconsin-Madison has 11 such communities. Learning community assignments in our institutional context are nonrandom, so we do not estimate their effects in this paper. Nonetheless, our findings corroborate the benefits of clustering students by academic skills on increased entrepreneurship.

Third, our research has important implications for gender diversity in entrepreneurship. Prior research shows that women are less likely to pursue entrepreneurship than men. (Greene, Hart, Gatewood, Brush, and Carter, 2003; Sørensen and Sharkey, 2014; Guzman and Kacperczyk, 2019). In our data, women have much lower baseline levels of entrepreneurial ideation, firm formation, and intentions. Effects of gender diversity on entrepreneurship are of particular interest as an opportunity to reduce the gender gap in entrepreneurship, and a substantial literature suggests that peer gender composition is an important determinant of outcomes in educational settings. We find evidence that peer gender diversity has statistically significant opposite effects on entrepreneurial intentions for women (negative) and men (positive), con-

tributing to this gap. This is consistent with existing work showing positive effects of girls and women in academic settings on the academic achievement and choices of their classmates, as measured by primary school test scores (Gottfried and Harven, 2015), college dropout postponement (Oosterbeek and Van Ewijk, 2014), and college majors (Zölitz and Feld, 2021). We also investigate the effect of dorm-level gender integration, with statistically insignificant results.

Fourth, to estimate effects of dorm and peer assignments, we extend the methods of Abdulkadiroğlu, Angrist, Narita, and Pathak (2017) that calculate and control for expected dorms and peers implied by a random room assignment mechanism that leverages both student room preferences and randomized tiebreaker lotteries to (quasi-randomly) assign students to rooms. These expected treatment controls are sufficient to eliminate omitted variables bias from unmodeled determinants of corresponding treatment assignments (Borusyak and Hull, 2020), allowing us to consistently estimate their effects. Unlike common applications of related methods that estimate effects of school assignments (Angrist, Hull, Pathak, and Walters, 2021; Angrist, Gray-Lobe, Idoux, and Pathak, 2022), we are unable to replicate the observed room assignments with a well-defined room assignment mechanism. Our insight is that the results from Borusyak and Hull (2020) and Abdulkadiroğlu, Angrist, Narita, and Pathak (2017) nonetheless hold for estimation of the effects of the *simulated* peer traits on both outcomes and realized peer traits, which permit us to use our misspecified simulated peer and dorm assignments as instruments for realized peer and dorm assignments. We find dramatically different results when we control for expected peer assignments using our method relative to simpler methods that do not control for nonrandom determinants of peer assignments.

In summary, our results suggest fruitful policy interventions relating to housing assignments that universities can implement to increase entrepreneurship rates and reduce the entrepreneurship gender gap. Extrapolations from our estimates suggest that assigning students to roommates with identical ACT verbal scores would increase entrepreneurial ideation by 26 percentage points from a baseline of 23 percent. Meanwhile, similar extrapolations suggest that completely eliminating neighbor-level gender-integration within floors would reduce the gender gap in entrepreneurial intentions by 15 percentage points from a baseline of 20 points, eliminating 75% of the gender gap in intentions at this very early point in the entrepreneurship pipeline. These extrapolations go outside of the variation we commonly find in our data, but nonetheless suggest substantial potential effects from implementation of policies based on our findings.

## 2. Data and Institutional Setting

### 2.1. Data Description

Our measures of students’ entrepreneurial proclivity come from a survey run from 2016 to 2019 of the undergraduate population of the University of Wisconsin-Madison, a large, public research university. The survey consists of fewer than ten questions which are designed to measure the career aspirations of the student population, including entrepreneurship. We merge our survey with university administrative data from both the University Registrar and University Housing. The administrative data contains information from 2016-2019 on the rooms students live in for years in which they live in residence halls, as well as a wealth of baseline variables such as gender, race, and intended major.<sup>1</sup> Importantly, the data from University Housing contains the residence hall preference lists provided by students, which is used to match them to their most-preferred dorms when possible.

Our empirical sample is formed of the subset of undergraduate students who live in dorm rooms during the fall term of their first year that are assigned the preceding summer and who respond to our survey on entrepreneurship. There are 24,265 undergraduate students who live in dorms in their first fall that are assigned during the preceding summer, of whom 6,919 (29%) also respond to our survey. We omit higher-year students from our sample, even if they live in the dorms, because they are relatively rare, we expect peer exposure to affect them less than first-year students, and because we observe many of them as first-years. Similarly, we omit first-year students who do not receive university residence hall room assignments the summer before their first fall term (for instance, due to missing the dorm preference deadline) because they are not included in the university’s general room allocation process that is central to our empirical strategy. We still make use of data on students who receive university residence hall room assignments the summer before the fall term who did not respond to our survey because we can observe their baseline characteristics and they are among the peers that potentially affect the students for whom we observe survey responses. Students who respond to our survey may do so at any point in their academic career, so we code an individual as having ideation if they ever report having ideas for products or services, and as having intention if they ever report intention to start a business, in order to maximize our effective sample size.

We note that the fundamental level of randomization is of students to dorm rooms, so

---

<sup>1</sup>Future versions of the paper will likely include additional years of administrative and survey data.

we think of all of our treatment variables as characteristics of dorm rooms. The dorm room characteristics we code relate to the baseline traits of other inhabitants of the room, baseline traits of inhabitants of nearby rooms, and traits of the residence hall in which the room is located. We have priors based on the extant literature that exposure to diverse peers will increase entrepreneurship, so we focus our attention on peer traits relevant to diversity.

The peer traits that we calculate are average peer gender distance, average peer STEM major distance, average peer ACT math distance, and average peer ACT verbal distance. For each of these, we calculate the absolute value distance of an individual’s own trait from the values of each individual in a given distance bin, then take the average. We calculate these values for each individual’s roommate(s), nearest neighbors, and next-nearest neighbors, obtaining three treatment variables for each trait. Students are coded to have a STEM major if their intended major on their college application matches a major from the ICE list of STEM-designated majors (ICE, 2016). Students are coded as male or female based on their self-reported gender to university administration. ACT scores come from administrative data, and we impute ACT scores for students with missing scores using AP test count, AP test score average, high school rank, high school class size, and high school GPA.

To estimate peer effects, it is necessary to identify students’ peers. Students’ roommates are readily identified as those individuals occupying the same room at the same time as a given student. To identify neighbors, we reference university blueprints and manually code x, y, and z coordinates for the doorway for each room on campus using drafting software. We then calculate Manhattan distances in inches between each room and every other room. This allows us to identify each room’s nearest 3 rooms and next-nearest 4 rooms, the inhabitants of which are treated as 2 relevant bins of neighbors for the student(s) in the room.<sup>2</sup>

In addition to effects of exposure to various types of peers at the room level, we are also interested in effects of dorm-level assignments. We are particularly interested in characteristics of residence halls that are manipulable by university administrators, as these are most policy-relevant at both the university we study and others. With this in mind, we estimate effects of dorms that are fully gender-integrated (men and women can live next door to one another). Additionally, we estimate effects of assignments to the residence hall that houses the StartUp learning community, a selective residential community that houses 64 students per year with

---

<sup>2</sup>This coding is intended to define nearest neighbors as individuals in the three rooms directly adjacent to or directly across the hall from an individual, and next nearest neighbors as the individuals adjacent to nearest neighbors.

interests in entrepreneurship. Placement into the community itself is done according to a nonrandom application process, but assignment to the residence hall in which the learning community is housed is random, and has the effect of placing students in close proximity to a concentrated mass of peers with a revealed interest in entrepreneurship.

Our empirical strategy controls explicitly for the expected values of treatments, so additional controls are not necessary for identification of effects of interest. We will nonetheless include controls in our preferred specifications in the interest of increasing statistical precision. We control for gender, race, a nontraditional student indicator, a first-generation college student indicator, ACT math score, ACT verbal score, and major fixed effects (2-digit Classification of Instructional Programs codes). Summary statistics for the variables we use are available in Table 1.

Table 1: Descriptive Statistics

	Overall		Survey Respondents	
	Mean (1)	SD (2)	Mean (3)	SD (4)
Intentions	0.275	0.446	0.275	0.446
Ideas	0.232	0.422	0.232	0.422
Female	0.531	0.499	0.565	0.496
Neighbor Gender Distance	0.384	0.349	0.357	0.351
STEM	0.435	0.496	0.446	0.497
Roommate STEM Distance	0.412	0.485	0.414	0.486
ACT Math	28.634	3.741	28.825	3.713
Roommate ACT Math Distance	3.644	2.857	3.665	2.859
ACT Verbal	57.749	8.944	57.903	8.911
Roommate ACT Verbal Distance	8.801	7.206	8.766	7.170
Nontraditional	0.016	0.125	0.019	0.136
In-State	0.556	0.497	0.607	0.488
First Generation	0.183	0.387	0.181	0.385
Asian	0.071	0.257	0.059	0.236
Black	0.018	0.132	0.015	0.121
Hispanic	0.056	0.230	0.044	0.204
White	0.707	0.455	0.733	0.442
Other Race	0.149	0.356	0.149	0.356
International	0.091	0.288	0.095	0.293
No Roommate	0.039	0.193	0.042	0.201
Random Room Assignment	0.691	0.462	0.670	0.470
Random Roommate Assignment	0.387	0.487	0.407	0.491
Observations	24265		6919	

*Notes:* Means and standard deviations for outcomes, peer traits, and controls for all first-year students in dorms, as well as students in dorms that respond to our survey on entrepreneurship.

## 2.2. Room Assignment Process

Our estimation of peer effects relies on randomness in the room assignment procedures used by University Housing. Over 90% of 1st year undergraduate students live in residence halls each year, along with many higher-year students. Students submit their residence hall applications in the summer prior to arrival on campus and are assigned to rooms by University Housing staff via a proprietary procedure. After students are notified of their initial assignments, they may request revisions at any time before or after the beginning of the school year. We observe the initial (often random) assignments as well as any (nonrandom) changes in room assignments over time.

A substantial number of initial room assignments are nonrandom. Per interviews with Housing staff, the following general procedure is used:

1. Allow the athletics department to assign student athletes to rooms.
2. Allow students admitted to Learning Communities to claim preferred rooms in their Learning Communities.
3. Allow returning students (2nd years and up) to claim preferred rooms.
4. Allow nonstandard students (incoming transfer students and students above age 20) to claim preferred rooms.
5. Assign genders to rooms to ensure that dorms have space proportional to incoming student gender shares.
6. Match remaining students to rooms according to a random room assignment mechanism.
7. Adjust assignments using judgment about optimal assignments.

The students assigned via steps 1-4 will not be randomly assigned to rooms, but they may be randomly assigned to peers, as the neighbors assigned to them may be assigned according to the random mechanism. In most cases, the students assigned according to steps 1-4 will not have randomly-assigned roommates. Similarly, the university allows incoming 1st-year students to designate a preferred roommate with whom they will share a room. These students will similarly not be randomly assigned to roommates, though they will still be randomly assigned to dorms unlike those assigned via steps 1-4. The shares of randomly and nonrandomly assigned students for rooms and roommates are shown in Table 1.



The random procedure used by Housing is proprietary, and Housing staff may use judgment in assigning students to dorms and rooms (Step 7).<sup>3</sup> With the caveat that Housing staff judgment dominates other considerations, Housing staff conveyed to us the following information regarding the dorm assignment mechanism:

1. Assign each student on campus a random number.
2. Assign each set of predetermined roommates the minimum value of their random numbers.
3. Order students in ascending order of random numbers.
4. Proceed through the list from top to bottom, assigning each student to a room in their favorite dorm if possible, otherwise their next favorite, etc, with each student's assignment being resolved before considering subsequent students' preferences.

The assignment mechanism described by Housing staff for dorm assignments corresponds to the Random Serial Dictator (RSD) assignment mechanism for randomly assigned students.<sup>4</sup>

The above details are sufficient to describe dorm placements. Housing does not have official guidance for its staff regarding room placements within dorms. This presents a challenge for us in replicating room (and therefore peer) assignments observed in the data. Empirically, we find that students clustered together in the random tie breaker order are also clustered geographically within dorms. We will discuss the methods we use to deal with this in detail in section 3.4..

### 3. Empirical Strategy

#### 3.1. Dorm Choice

We estimate effects of peer gender distance, peer STEM distance, ACT score distance, dorm gender integration, and dorm entrepreneurial culture on entrepreneurial ideation and intention. Raw comparisons between individuals with different peers or dorms will not identify treatment effects for peer or dorm traits of interest if individuals choose dorms or rooms based on their

---

<sup>3</sup>For instance, in interviews Housing staff reported attempting to place students from the same high school in different rooms, with the intuition that those individuals would have explicitly requested each other as roommates if they had desired such a match.

<sup>4</sup>Random Serial Dictator is a special case of the Deferred Acceptance mechanism. A more comprehensive description of the room assignment mechanism described is that Deferred Acceptance is used, where rooms give individual-specific priority (unobserved to the researcher) to particular athletes, non-1st-years, nontraditional students, and students admitted to learning communities, with those students have corresponding preferences for their rooms.

entrepreneurial proclivity or unobserved personality traits correlated with entrepreneurial proclivity. For instance, if students who are more open to experience are more entrepreneurial and also tend to prefer dorms with diverse student populations, we may observe correlations between exposure to diversity and entrepreneurship that will not reflect peer treatment effects.

To address selection bias, our empirical strategy leverages the randomness in the room assignment mechanism used by University Housing. To facilitate the discussion, we adapt notation from Abdulkadiroğlu, Angrist, Narita, and Pathak (2017) to describe room assignments. A room assignment mechanism assigns individuals  $i = 1, 2, \dots, N$  in set  $\mathcal{I}$  to room-spots (beds) indexed by  $s$ , with  $s = 1, 2, \dots, S$ . There are  $N$  total students and  $S$  total room spots, where a particular feature of our setting is that  $N = S$ .

Each student  $i$  in our setting submits a complete preference ordering over dorms,  $\succ_i$ , where  $A \succ_i B$  means that student  $i$  prefers dorm  $A$  to dorm  $B$ .<sup>5</sup> In principle, we allow for each room  $s$  to give priority  $\rho_{is}$  to each student  $i$  with  $\rho_{is} < \rho_{ij}$  indicating the room prefers to house  $i$  over  $j$ . We do not have details on the priorities that rooms place on students ex ante, but describing rooms as giving students de facto priorities allows us to describe unobserved-to-the-researcher nonrandom determinants of housing assignments with standard allocation mechanism language and notation. For instance, we describe rooms that house women/men as giving priority to their respective gender, and we describe rooms that house non-1st-year-students as giving priority to the individual student who resides in the room, with all other rooms giving priority to other students.

Each student has a type indicated by  $\theta_i = (\succ_i, \boldsymbol{\rho}_i)$ , where  $\boldsymbol{\rho}_i$  is a vector that contains  $i$ 's priority in each room. In our application, we observe  $\succ_i$  for all students, but we do not necessarily observe  $\boldsymbol{\rho}_i$ , as this incorporates the judgment used by Housing staff in allocating students to rooms in accordance with University objectives. Each student is also assigned a random tie-breaker  $r_i$  which we also observe. A mechanism  $\varphi$  is a mapping from students' types and random numbers into room-spot assignments for all students. A student's room-spot assignment as determined by mechanism  $\varphi$  is given by  $\mu_i^\varphi(\theta, r)$ , with  $i$ 's realized assignment  $\mu_i(\theta, r) = \mu_i^*(\theta, r)$  being defined as the assignment determined by the true (unknown) allocation mechanism, denoted by  $\varphi = *$ .<sup>6</sup>

---

<sup>5</sup>Students are allowed to list up to 25 dorm preferences. We assume that students who do not fill their preference ranking list are indifferent between omitted dorms.

<sup>6</sup>The  $\mu^\varphi$  notation allows us to readily express assignments under alternative mechanisms. For instance,  $i$ 's room-spot assignment under the Boston dorm assignment mechanism, followed by a room assignment mechanism that places students in rooms in ascending room number order would be denoted  $\mu_i^{Boston, RN^\uparrow}(\theta, r)$ .

The room choice problem we describe assigns students to rooms, but it is impractical to conceive of room-level treatments. First, very few (usually two) students are assigned to each room at a given time, so we will be woefully underpowered to identify treatment effects for rooms. Second, any room-level treatment effects will lack policy-relevance outside of our immediate institutional context. Third, the primary treatments of interest to our study are the traits of peers, which are not fixed at the room level (with the notable exception of gender), but are rather characteristics of the other individuals inhabiting them. We therefore describe individual  $i$ 's treatment status  $D_{ik}(\mu^\varphi(\theta, r))$  for treatment  $k = 1, \dots, K$  as a function of the vector of room-spot assignments under mechanism  $\varphi$  for themselves (dorm effects) and other students (peer effects).

### 3.2. Treatment Effect Estimation via Stratification

The empirical strategies we implement leverage conditional randomization in the room assignment procedure. A general representation of the models we will estimate (omitting  $k$  subscripts to distinguish peer traits) is

$$Y_i = \beta_0^\varphi + \beta_D^\varphi D_i^\varphi + \beta^\varphi \mathbb{E}[D_i^\varphi | \theta] + \epsilon_i^\varphi, \quad (1)$$

where  $Y_i$  is a binary indicator for whether a student has an idea for a product or service or intention to start a company,  $D_i^\varphi = D_{ik}(\mu^\varphi(\theta, r))$  is the value of a treatment such as the share of  $i$ 's roommates under a given mechanism with different major STEM status,  $\mathbb{E}[D_i^\varphi | \theta]$  is the expected value of this treatment variable for student  $i$ , and  $\epsilon_i^\varphi$  captures unobserved determinants of outcomes.  $\beta_0^\varphi$  is a constant and the coefficient  $\beta_D^\varphi$  is the effect of treatment assignment as implied by mechanism  $\varphi$ . Conditioning on the expected value of treatment (the propensity score for binary treatments) eliminates selection bias by siphoning out the variation in treatment that is nonrandom (Rosenbaum and Rubin, 1983; Imbens, 2000; Borusyak and Hull, 2020). We model treatments as defined by an arbitrary mechanism  $\varphi$ , but our ultimate goal is estimation of effects of realized treatment,  $D_i$ .

The key to identification of the effect of  $D_i^\varphi$  is to properly calculate  $\mathbb{E}[D_i^\varphi | \theta]$ , or to include covariates that are sufficient statistics for it. Controlling for the expected value of treatment is fundamentally similar to controlling for all nonrandom determinants of treatments,  $\theta_i$  (Rosenbaum and Rubin, 1983). While we will rely on simulations to directly calculate  $\mathbb{E}[D_i | \theta]$ , our

identification strategy is closely related to the case in which students are assigned randomly to rooms conditional on observed covariates such as gender and dorm. More generally, our empirical strategy is a stratified randomized research design with selection on observables.<sup>7</sup>

There are three noteworthy approaches for estimating peer effects that will not reliably provide consistent estimates in our setting. The first is to condition on dorm assignments under the assumption that students are randomly assigned to rooms within their endogenously chosen dorm, such that the dorm chosen by student  $i$  is modeled as an element of  $\theta_i$ , but their specific room is not. We are concerned that students placed into dorms early in the assignment procedure will both rank their dorm highly and be placed near to one another within the dorm — introducing nonrandom dorm-preference clustering within dorms.<sup>8</sup> The next approach is to control for every possible permutation of dorm preferences. There are 21 dorms on campus and there is no limit to the number of dorms a student can list, so in principle this would require controlling for fixed effects for  $21!$  permutations of dorm preferences, likely interacted with student gender (with exceptions for permutations that describe zero students). Not only would this approach condition out the vast majority of treatment variation in the data, but it also may not even be sufficient to control for students' types, as Housing staff may use information other than preferences and gender to make room assignments ( $\theta$  contains more than just preferences and gender). The final approach is to calculate  $\mathbb{E}[D_i^*]$  using details of the room allocation mechanism used by Housing either via simulations or via analytic formulas for the expected values of treatment, as provided by Abdulkadiroğlu, Angrist, Narita, and Pathak (2017). In our case, we do not know the mechanism used by Housing, so we cannot calculate  $\mathbb{E}[D_i^*]$ .

While we cannot calculate  $\mathbb{E}[D_i^*]$ , we can calculate  $\mathbb{E}[D_i^\varphi]$  for an arbitrary mechanism,  $\varphi$ , that we specify. As explained by Abdulkadiroğlu, Angrist, Narita, and Pathak (2017), treatment effect estimates can be obtained for conditionally random assignment to treatment as long as the assignment mechanism satisfies the Equal Treatment of Equals (ETE) condition. The ETE condition requires that students with the same type (preferences,  $\succ$ , and priorities,

---

<sup>7</sup>In an institutional setting where students pick their dorms, but rooms and roommates were assigned randomly conditional on dorm, the preferred dorm would be included in  $\theta_i$  and expected peer exposure could be calculated via simulation as we do in this paper. Such cases have a relatively small number of strata (the number of dorms, or perhaps double this to account for gender-specific assignments), making calculation of expected treatments unnecessary.

<sup>8</sup>Our concern is validated by the relationship between peer tiebreakers and our simulated peer tiebreakers (which are generated using a simulation that produces extreme clustering) shown in Table 2. A particularly extreme form of this sort of clustering would occur if Housing staff (for instance) filled dorms in room number order, with low room numbers filled exclusively with students with low/lucky priority numbers who rank that dorm 1st in their preference list.

$\rho$ ), have identical assignment probabilities for each room. Following the argument of Abdulkadiroğlu, Angrist, Narita, and Pathak (2017), we can estimate the effects of treatment  $D^\varphi$  for any mechanism that satisfies ETE by controlling for  $\mathbb{E}[D_i^\varphi]$ , which we can calculate via simulation.<sup>9</sup>

It follows from the above that  $\beta^\varphi$ , the effect of the treatment agents would get were they assigned according to mechanism  $\varphi$ , is identified. This parameter does not have immediate policy relevance, despite its causal interpretation, unless it happens to correspond to realized treatment assignments such that  $D_i^\varphi = D_i^*$ . However, following the same arguments as we give above for the effect of  $D_i^\varphi$  on  $Y_i$ , we can also estimate the effect of  $D_i^\varphi$  on  $D_i$  from

$$D_i = \gamma_0^\varphi + \gamma_D^\varphi D_i^\varphi + \gamma^\varphi \mathbb{E}[D_i^\varphi | \theta] + u_i^\varphi, \quad (2)$$

where  $\gamma_D^\varphi$  is identified as the effect of assignments as implied by mechanism  $\varphi$  on actual assignments  $D_i$  under the same argument made above for its effect on  $Y_i$ . The equations (1) and (2) together form the reduced form and first stage, respectively, of a just-identified instrumental variables model. It follows from the identification of  $\beta_D^\varphi$  and  $\gamma_D^\varphi$  that the effect of actual assignments on outcomes is also identified as

$$\beta_D^* = \frac{\beta_D^\varphi}{\gamma_D^\varphi}$$

under standard instrumental variables assumptions, which we will discuss presently.

### 3.3. Identification Assumptions

The instrumental variables independence assumption requires that any systematic relationship between students' treatment assignments and the value of the instrument are captured by control variables. Because we explicitly control for the expected value of the instrument for each individual, this condition is almost sure to be satisfied (Imbens, 2000; Hirano and Imbens, 2004; Borusyak and Hull, 2020). The threat to this condition is overfitting, in which the mechanism that we specify accidentally explains observed treatments well without distinguishing between types of students with respect to expected treatments. We will discuss the threat to independence from overfitting in greater detail in section 3.4..

The instrumental variables exclusion restriction requires that treatment assignments as de-

---

<sup>9</sup>It seems to us that the ETE condition is an assumption on the knowledge of the researcher of the details of a mechanism and observability by a researcher of its inputs  $\theta$  and  $\rho$ . From the perspective of the mechanism or the individuals implementing it, it is tautological that equals are treated equally.

terminated by a chosen mechanism only affect outcomes via the actual treatment. The possibility that simulated treatment assignments affect outcomes through a channel unrelated to Housing assignments, such as students assigned by the mechanism to the StartUp dorm happening to have entrepreneurial parents, is addressed by controlling for expected values of instruments.

A more subtle issue relating to the exclusion restriction is that an infinite number of unmodeled peer and dorm characteristics are contained within the error term  $\epsilon^{\varphi}$  in equation (1), and these unmodeled peer and dorm traits are certainly correlated with treatments of interest to our study. For instance, assignment to the StartUp dorm also means being assigned across the street from the Business School, being assigned to the most popular dorm on campus (as measured by the percentage of students listing it as their first preference), being assigned in close proximity to the disability resources center, and being assigned in close proximity to a recreation center. Some, but not all, of these traits are shared by other dorms, and the intersection of these traits uniquely defines the StartUp dorm. It follows that were we to include these alternative traits in our empirical specifications with sufficient flexibility, they would be perfectly collinear with the StartUp dorm treatment assignment, rendering us unable to distinguish their effects. We address this issue by modelling treatments that we believe are most salient to our outcomes *ex ante*, with the caveat that further study in other settings would be particularly useful for distinguishing between multiple explanations for effects we find.<sup>10</sup>

The monotonicity assumption (Imbens and Angrist, 1994) implies that instruments that increase the likelihood of treatment on average do so for everyone. In our context, this means that receiving a particular type of simulated peer never makes any individuals in the data less likely to receive that type actual peer. We consider two cases.

The first is the case in which our room assignment mechanism is fundamentally misspecified, in which case it will fail to predict actual assignments at random for some individuals. This may create situations in which an individual with a particular type of simulated peer does not have the same type of actual peer. Problematically, it could also be the case (as with random allocation mistakes) that such an individual would have the same type of actual peer if they did not have such a simulated peer. The existence of any such “defiers”, which we deem very likely, constitutes a violation of Imbens and Angrist’s monotonicity condition. However,

---

<sup>10</sup>This issue is paralleled in the school choice literature. For instance, the student assignments to charter schools described in Abdulkadiroğlu, Angrist, Narita, and Pathak (2017) are likely correlated with innumerable potentially-measurable pedagogical or administrative practices at those schools. If the true drivers of student outcomes are these practices, rather than a school’s status as a charter, the results from AANP will only have external validity to other contexts with similar correlation between charter school status and these practices.

because these individuals' assignments are simulated incorrectly at random because of our misspecification of the assignment mechanism, it is implausible that they are systematically different from other students with respect to the effects of peer assignments. We therefore invoke the weaker compliers-defiers condition described by De Chaisemartin (2017), which implies that we still recover meaningful local average treatment effects when a sufficiently large subset of the compliers share treatment effects with the defiers.

A more worrying possibility is that there are cases when our chosen room assignment mechanism matches the original room assignments made by housing staff in the summer, but students request reassignments. This mirrors monotonicity issues that occur in encouragement designs, such as when financial inducements may crowd out intrinsic motivation (Gneezy and Rustichini, 2000). Students (or their parents) who find a way to achieve particular dorm or peer assignments regardless of their initial placements are always takers or never takers in the language of Angrist, Imbens, and Rubin (1996), and will not contribute to estimates. However, students who tend to request reassignments regardless of their initial assignments, for instance because they have unrealistic expectations about how well they will get along with their initially assigned roommate, will be defiers. It seems reasonable that these individuals may have somewhat different responsiveness to peer exposure on average than compliers, threatening the validity of the De Chaisemartin (2017) compliers-defiers condition. We explore the likely extent of this issue in Appendix Appendix B:, finding that room switches in the summer are quite rare which gives us confidence that sufficient compliers exist that cancel out any such defiers.

The stable unit treatment value assumption (SUTVA) implies that there is no unmodeled treatment heterogeneity (Rubin, 1978, 1980, 1990). A commonly-discussed case of this is unmodeled spillovers, which is a particular concern in our setting. This assumption is violated if traits of peers in dorms that we do not include in our models (those of peers who are further than 7 rooms away from an individual) affect individuals directly or indirectly (by affecting their nearby peers). To distill this into a concrete example, it may be that some individuals have STEM peers who have STEM peers, while other individuals have STEM peers who have a mix of STEM and non-STEM peers, and these are fundamentally different types of peers with respect to their effects on outcomes.

In our view, a simpler example of a SUTVA violation is that of essential treatment heterogeneity, where the treatment assignment means different things for different individuals. Continuing with the STEM example, SUTVA will be violated if being assigned a Biology major

roommate has different effects than being assigned an Engineering major roommate, as we code both of these as STEM roommates. Because of both of these concerns regarding SUTVA in our application, we instead assume uniform unordered monotonicity (Harris, 2022), which imposes that having simulated peers of a particular type (STEM roommate) weakly increases assignment of all unobserved component treatment types (all types of STEM roommate). In simple terms, this assumption requires that instrumental variables do not cause within-treatment switching, such as if being assigned a simulated STEM roommate increased the likelihood of having an Engineering major roommate and decreased the likelihood of a Biology major roommate.

In summary, our empirical strategy is especially robust to the independence assumption. Its robustness to the exclusion assumption relies on us choosing the most relevant treatments in a strict sense, but it relies only on the correlations between the treatments we choose and unobserved treatments in our sample resembling those in the population to preserve policy-relevance of treatment effect estimands. We expect our instruments to satisfy complier-defiers monotonicity but to violate SUTVA, so we impose the alternative assumption of uniform unordered monotonicity which allows us to interpret the effects of modeled peer assignments as convex averages of unmodeled peer assignments.

### 3.4. Instrument Selection

In the prior section we discussed independence, exclusion, monotonicity, and SUTVA for our instrumental variables, but we did not discuss relevance - the requirement that instruments are predictive of treatments of interest. This section describes our procedure for choosing a room assignment mechanism to use to construct IVs with the intention of avoiding weak instruments and maximizing our power for estimating treatment effects.

We begin the discussion with a general representation of the problem, which will establish a connection to the model selection literature. The key insight is that there are an infinite number of potential mechanisms that we could propose for assigning students to dorm rooms. It follows that there are an infinite number of valid simulated instruments for peer and dorm assignments. These realizations lead us to instrumental variables model selection methods, such as those described by Belloni, Chen, Chernozhukov, and Hansen (2012) and Belloni, Chernozhukov, and Hansen (2014). Broadly speaking, these methods consider first stage equations that are similar to equation (2) that in principle allow for an arbitrarily large number of instrumental variables,



such as

$$D_i = \gamma_0 + \sum_{\varphi=1}^{\Phi} (\gamma_D^{\varphi} D_i^{\varphi} + \gamma_E^{\varphi} \mathbb{E}[D_i^{\varphi} | \theta]) + u_i, \quad (3)$$

with  $\varphi = 1, 2, \dots, \Phi$  indexing the proposed room assignment mechanisms, with the important point being that  $\Phi \gg N$ .

Model selection methods search over many specifications like those in equation (3) and identify the one(s) that include the strongest instruments. While our problem is the same on a fundamental level, we face nonstandard challenges relative to common applications that use these methods for instrument selection. First, for us to construct an instrument, we must conceive of (or identify in the literature) a room assignment mechanism, code the mechanism in statistical software, and run the code that generates instruments using the mechanism. This is costly in terms of cognition, human time, and computational time. Secondly, restricting ourselves to a single best-performing mechanism is particularly attractive for interpreting effects in the context of our institutional setting.<sup>11</sup> Finally, as we will discuss further in a moment, we are at particular risk of overfitting equation (3) if we search over an arbitrary number of mechanisms.

The particular risk of overfitting in our application comes from each student receiving a unique value of the tiebreaker,  $r_i$ , with simulated room-spot assignments being a unique and deterministic function of  $r$  and observables. The implication of this is that it is possible to construct simulated instruments and expected simulated treatments that (1) satisfy Equal Treatment of Equals, (2) perfectly rationalize all observed assignments, yielding a deterministic first stage, and either (3a) explain all of the variation in actual treatments with simulated treatments, generating extremely strong instruments or (3b) explain all of the variation in actual treatments with expected simulated treatments, generating extremely weak instruments.

As an example of case (3a) above, consider the egregiously overfit mechanism  $\varphi = **$  that satisfies  $\mu = \mu^*(\theta, r) = \mu^{**}(r)$ . This mechanism effectively observes the room-spot student  $i$  was assigned and infers that any student assigned random tie-breaker  $r_i$  in a counterfactual assignment allocation would receive room-spot  $\mu_i$ , such that  $\mu_j^{**}(r_i) = \mu_i^{**}(r_i)$  for all  $j \neq i$ . This mechanism trivially satisfies ETE because all students have equal probabilities of all room

---

<sup>11</sup>In other words, using simulated instruments from a single mechanism that is established in the mechanism design literature allows us to describe our reduced form effects from equations like (1) as “the effect of peers a person is assigned under mechanism  $\varphi$  on outcome  $Y_i$ ”. The inclusion of instruments from multiple mechanisms weakens this intuition in our view.

assignments, implying that all students of the same type do as well. Because  $\mu_i(\theta, r) = \mu_i^{**}(r)$  for all  $i$ , it follows that the first stage in (2) collapses to

$$\begin{aligned} D_i &= \gamma_0^{**} + \gamma_D^{**} D_i^{**} + \gamma^{**} \mathbb{E}[D_i^{**} | \theta] + u_i^{**} \\ &= D_i^{**}. \end{aligned} \tag{4}$$

$D_i^{**}$  is a strong instrument for  $D_i$  (its F-statistic is unboundedly large regardless of sample size), with the expected value of simulated treatments playing no role in predicting actual treatment assignments. There is minimal variation in  $\mathbb{E}[D_i^{**} | \theta]$ , so it will fail to capture unobserved determinants of room assignments in (1) as well. It follows that implementing a two-stage least squares regression using  $D_i^{**}$  as an instrument for  $D_i$  while controlling for  $\mathbb{E}[D_i^{**} | \theta]$  is approximately equivalent to the naive OLS specification

$$Y_i = \beta_0 + \beta_D D_i + e_i, \tag{5}$$

which is unlikely to identify causal estimates due to selection, as discussed above.<sup>12</sup>

As an example of case (3b) above where a deterministic mechanism produces weak instruments, consider the deterministic-on-observables mechanism  $\varphi = \theta$  that satisfies  $\mu = \mu^*(\theta, r) = \mu^\theta(\theta)$ , where sufficient individual characteristics are contained in  $\theta$  to perfectly determine assignments. This mechanism effectively observes the room spot  $i$  was assigned and infers that they must have been nonrandomly assigned to that spot based on spot priorities over student characteristics, the simplest case of which is that preferences are ignored, and each room spots give special priority to the student that inhabits it. It follows that each individual would always be assigned their actual room in any counterfactual assignment allocation, with  $\mu_i^\theta(\theta, r) = \mu_i^\theta(\theta, r')$  for all  $i$  for any alternative set of tiebreakers  $r'$ . In this case, the first stage in (2) collapses to

$$\begin{aligned} D_i &= \gamma_0^\theta + \gamma_D^\theta D_i^\theta + \gamma^\theta \mathbb{E}[D_i^\theta | \theta] + u_i^\theta \\ &= \mathbb{E}[D_i^\theta | \theta]. \end{aligned} \tag{6}$$

Here,  $D_i^\theta$  is a weak instrument (with an F-statistic of zero regardless of sample size), with expected treatments completely explaining realized treatments.  $D_i^\theta$  is perfectly collinear with

---

<sup>12</sup>We say that instrumenting for  $D_i$  with  $D_i^{**}$  while controlling for  $\mathbb{E}[D_i^{**} | \theta]$  is approximately equivalent to OLS with no controls because  $\mathbb{E}[D_i^{**} | \theta]$  still accounts for individuals' inabilities to be their own peers. Each individual's value of  $\mathbb{E}[D_i^{**} | \theta]$  will be constructed as the sample leave-one-out mean of  $D$ , which is similar to an OLS regression controlling only for individual  $i$ 's own trait with no other room assignment strata indicators.

$\mathbb{E}[D_i^\theta|\theta]$  so treatment effects in (1) are unidentified. We prefer mechanism  $\theta$  to mechanism  $**$  because it honestly reports its usefulness with an F-stat of zero, but both mechanisms are inadequate for treatment effect estimation.

In order to identify a mechanism that avoids overfitting while also predicting treatment assignments, we place constraints on ourselves in our mechanism search. First, we interviewed University Housing prior to attempting to rationalize observed assignments with any mechanism, and we constrain ourselves to dorm assignment mechanisms that we discussed with them.<sup>13</sup> Second, we require that all mechanism guesses make use only of characteristics that Housing told us they use: random tiebreakers, gender, nontraditional student status, year of study, athlete status, learning community status, predetermined roommate status, and dorm preferences. Third, we restricted ourselves to monotonic functions of random tiebreakers; that is, we require that assignment to preferred dorms are monotonically decreasing in the value of the random tiebreaker (early ranked students get their preferences over late students). Fourth, we do not explicitly target our peer or dorm treatments of interest, but instead we target all individual dorm assignments and peer random tiebreaker assignments. In other words, we do not attempt to find a mechanism that rationalizes assignment to coed dorms or to peer gender diversity, we instead attempt to find a mechanism that rationalizes realized dorm assignments generally and realized peer tiebreaker values.<sup>14</sup>

We consider three dorm assignment mechanisms and three room assignment mechanisms that condition on dorms being assigned according to the dorm mechanism that best rationalizes realized dorm assignments. All of the mechanisms we consider have some common components. First, we hold fixed rooms for individuals described by Housing as not being randomly assigned: athletes, nontraditional students, students in learning communities, and students with no recorded random tiebreaker. These individuals will not contribute to dorm effect estimates because their expected dorm is equal to their realized dorm. They will contribute to peer effect estimates because though they are not randomly assigned to rooms, their roommates and neighbors may be. Second, following advice from Housing, we place students in remaining dorms or room spots at random after all other students are assigned if their preferences are insufficient to place them according to the other rules of a mechanism. This is primarily relevant for

---

<sup>13</sup>In addition to constraining us in terms of specification search, this also drastically reduced our workload.

<sup>14</sup>If Housing staff place individuals sequentially in a floor or hall who are sequential in the random tiebreaker order, we will see substantial geographic clustering with respect to random tiebreakers and our guesses that replicate this clustering will also replicate peer assignments of interest to the study.

individuals who do not provide complete preferences on their dorm preference sheet.

We consider three dorm assignment mechanisms. First, we consider random serial dictator, which assigns each individual to their preferred dorm if it is available, then to their next preferred, and so on, in order  $r_i = 1, 2, \dots, N$  without reference to other individuals. Next, we consider the Boston mechanism, which assigns individuals in order  $r_i = 1, 2, \dots, R$  to their 1st preferred dorm if possible, otherwise skipping them. Then it repeats this for unassigned individuals for their 2nd preferred dorm, and so on, until all students are assigned. Finally, because Housing reported occasionally deviating from RSD in an effort to fill unpopular dorms with students who like them *relatively* well, we consider a mechanism we call Boston→RSD which assigns individuals in order  $r_i = 1, 2, \dots, R$  to their 1st preferred dorm *if and only if it is on a list of unpopular dorms* if possible, otherwise skipping them. Then it repeats this for unassigned individuals for their 2nd preferred dorm, and so on, until all unpopular dorms are filled. It then places remaining students in dorms according to RSD.

We also consider three room assignment mechanisms, conditional on dorms being assigned according to the dorm mechanism that best rationalizes realized dorm assignments. The first, which we term Room ID order, is that students are placed in rooms in ascending room ID order as they are placed into a hall, where room ID is an administrative record that is distinct from room numbers. The second, which we term Room # order, is that students are placed in rooms in ascending room number order as they are placed into a hall, where room number is the publicly observable number for each room, viewable while standing outside the door. Finally we consider a Geographic order, which places students into rooms in a zig-zag using x and y coordinates for room doorways, and ascending z (floor) order.<sup>15</sup>

Standard tests of instrument strength would estimate equations such as (3) and consider the F-statistic associated simulated instruments from each mechanism. We opt for an alternative approach due to the computational intensity of calculating  $\mathbb{E}[D_{ik}(\mu^\varphi(\theta, r))|\theta]$  for all  $i$  and  $k$  for each mechanism. We estimate the following seemingly unrelated regression (SUR) model to evaluate mechanism accuracy for all  $J$  dorms in  $j = 1, 2, \dots, J$  for each dorm assignment

---

<sup>15</sup>There are myriad plausible ways to code geographic room orders, all of which will produce very similar peer assignments if they respect our monotonic tiebreaker constraint. Generally, all rotations and reflections of assignments for symmetric dorms will produce identical peer assignments, while less extreme diversions (or approximations of such diversions in asymmetrical dorms) will produce similar peer assignments. We consider only a single geographic order rather than embarking on a specification hunt among many extremely similar mechanisms that explicitly replicate tiebreaker clustering geographically with room assignment clustering.

mechanism,

$$\begin{aligned}
Dorm_{i,1} &= \alpha_{Dorm}^{\varphi} Dorm_{i,1}^{\varphi} + u_{i,1} \\
Dorm_{i,2} &= \alpha_{Dorm}^{\varphi} Dorm_{i,2}^{\varphi} + u_{i,2} \\
&\vdots \\
Dorm_{i,J} &= \alpha_{Dorm}^{\varphi} Dorm_{i,J}^{\varphi} + u_{i,J},
\end{aligned} \tag{7}$$

where we emphasize that the constant is constrained to 0 and all equations in the model are constrained to share a single slope coefficient.  $\alpha_{Dorm}^{\varphi}$  has the attractive property of giving the probability of individual  $i$  having a realized assignment to dorm  $j$  conditional on having simulated assignment to dorm  $j$ . We similarly estimate accuracy measures for peer tiebreakers as

$$\begin{aligned}
r_{i,0} &= \alpha_r^{\varphi} r_{i,0}^{\varphi} + \delta_{i,0}^{\varphi} + u_{i,0} \\
r_{i,1-3} &= \alpha_r^{\varphi} r_{i,1-3}^{\varphi} + \delta_{i,1-3}^{\varphi} + u_{i,1-3} \\
r_{i,4-7} &= \alpha_r^{\varphi} r_{i,4-7}^{\varphi} + \delta_{i,4-7}^{\varphi} + u_{i,4-7},
\end{aligned} \tag{8}$$

where  $r_{i,0}$  gives the average tiebreaker for individual  $i$ 's realized roommate,  $r_{i,1-3}$  gives the same for next door neighbors,  $r_{i,4-7}$  gives the same for the next nearest neighbors. The  $\varphi$  superscript gives the same values for simulated assignments, with  $\delta_i^{\varphi}$  denoting simulated dorm fixed effects. We report dorm assignment mechanism accuracy,  $\alpha_{Dorm}^{\varphi}$ , in Panel 1 of Table 2 and room assignment mechanism accuracy,  $\alpha_r^{\varphi}$ , in Panel 2 of Table 2 for all the mechanisms we consider.

In the interest of thoroughness, we cross-validate mechanism accuracy for 5 samples of students in our data who are randomly assigned to dorms and rooms (where we keep only those with roommates for the room assignment accuracy). The Random 1 and Random 2 samples are randomly chosen mutually exclusive halves of all randomly assigned students. We also consider 2016 (a year when University Housing staff reported they exercised less judgment in room assignments) separately from later years (when Housing staff exercised more judgment in room assignments). Finally, we also calculate accuracy for the entire sample of randomly-assigned students. We choose mechanisms with the highest accuracy measures, RSD and Room #, for the total sample to construct instruments for inclusion in 2SLS models described by equations (2) and (1), but we are encouraged by the cross-sample consistency in accuracy-maximizing

mechanisms.

Table 2: Mechanism Selection

	Random 1 (1)	Random 2 (2)	2016 (3)	2017-2019 (4)	Total (5)
Panel 1: Dorm Assignments					
RSD	0.885 (0.002)	0.872 (0.002)	0.958 (0.002)	0.859 (0.002)	0.879 (0.002)
Boston	0.836 (0.003)	0.779 (0.003)	0.865 (0.004)	0.813 (0.003)	0.815 (0.002)
Boston $\rightarrow$ RSD	0.000 (0.000)	0.628 (0.004)	0.740 (0.005)	0.670 (0.003)	0.676 (0.003)
Observations	2319	2319	1437	3201	4638
Panel 2: Peer Assignments					
Room ID	0.493 (0.011)	0.505 (0.011)	0.452 (0.013)	0.483 (0.010)	0.498 (0.008)
Room #	0.519 (0.011)	0.502 (0.011)	0.464 (0.013)	0.490 (0.010)	0.509 (0.008)
Geographic	0.526 (0.011)	0.490 (0.011)	0.452 (0.013)	0.494 (0.010)	0.507 (0.008)
Observations	2091	2113	1339	2865	4204

*Notes:* Mechanism accuracy for dorms as measured by  $\alpha_{Dorm}^{\varphi}$  from the seemingly unrelated regression model (7) in Panel 1 and mechanism accuracy for rooms as measured by  $\alpha_r^{\varphi}$  from the seemingly unrelated regression model (8) in Panel 2. Mechanism details are described in the text. Robust standard errors of accuracy statistics in parentheses.

Our selected mechanisms produce simulated and expected dorm and peer treatment assignment statuses described in Table 3. There are three main takeaways from this table. First, our simulated and expected treatments have means that are extremely close to the means of realized treatments. Second, realized and simulated treatments have significantly more variance than expected treatments. Our identification strategy uses residual variation that is common to both simulated and realized treatments after conditioning on expected treatments to identify effects, so relatively low variance in expected treatments is good news.<sup>16</sup> Finally, we do much better at matching dorm assignments than peer assignments, which is unsurprising given Housing staff’s description of having fixed heuristics for dorm assignments but not room assignments.

<sup>16</sup>Note that expected treatments vary well over half as much as actual treatments, suggesting that most of the variation in peer and dorm assignments are driven by selection.

Table 3: Treatment Assignment Description

	Survey Respondents	
	Mean (1)	SD (2)
Neighbor Gender Distance	0.357	0.351
Simulated Neighbor Gender Distance	0.361	0.354
Expected Neighbor Gender Distance	0.362	0.270
Roommate STEM Distance	0.414	0.486
Simulated Roommate STEM Distance	0.424	0.484
Expected Roommate STEM Distance	0.425	0.382
Roommate ACT Math Distance	3.665	2.859
Simulated Roommate ACT Math Distance	3.697	2.829
Expected Roommate ACT Math Distance	3.710	2.367
Roommate ACT Verbal Distance	8.766	7.170
Simulated Roommate ACT Verbal Distance	8.967	7.233
Expected Roommate ACT Verbal Distance	8.939	5.997
StartUp Dorm	0.148	0.355
Simulated StartUp Dorm	0.148	0.355
Expected StartUp Dorm	0.153	0.258
Coed Dorm	0.678	0.467
Simulated Coed Dorm	0.681	0.466
Expected Coed Dorm	0.685	0.379
No Roommate	0.042	0.201
No Simulated Roommate	0.047	0.213
Observations	6919	

*Notes:* Means and standard deviations for outcomes, peer traits, and controls for all first year students in dorms as well as those that respond to our survey on entrepreneurship.

## 4. Results

We estimate effects of dorm gender integration, dorm entrepreneurial culture, peer gender diversity, and peer academic diversity on entrepreneurial intentions and ideas using two stage least squares specifications defined by equations (1) and (2). It is possible in principle for us to include all treatments in a single specification, but we instead make use of specifications only included closely linked treatments in the interest of obtaining administratively actionable results and to preserve instrument strength.

Correlated treatment assignments produce particular complications for generating administratively feasible actionable insights for models containing multiple treatments. For instance, we find that verbal skill diversity has negative effects on ideation. It possible that part of this total effect is driven by correlation between verbal skill diversity and gender diversity (which has statistically insignificant effects in the same direction). For a complicated room assignment policy that conditions jointly on many student characteristics to make optimal room assignments, results from specifications including the full set of treatments are relevant. For simpler policies based on effects of a single peer trait, single-treatment specifications answer the policy-relevant question. In other words, it is policy relevant to identify the total (statistically significant) effect of verbal skill diversity on ideation even if part of the effect of is actually due to gender diversity.

In addition to our interest in obtaining administratively actionable results, we also want to obtain statistically sensible results. In our specifications, including a large number of treatments necessitates also including a large number of expected treatment controls. Even though each treatment's first stage only requires its own expected treatment as a control, standard 2SLS also includes expected treatments for all other treatments in the first stage. This has the effect, in small samples, of consuming residual variation in the first stage that could be attributable to instruments with control variables, resulting in weak instrument problems.<sup>17</sup> With this in mind we focus on parsimonious models with small numbers of closely related treatment variables.

We begin our discussion of estimates with dorm effects, as these are the treatment assignments that we do the best job of predicting (see Table 2). Results for the effects of gender-

---

<sup>17</sup>A potential answer to this issue is to impose exclusion restrictions on simulated and expected treatments other than those that correspond to a given realized treatment for each equation in the first stage. This involves excluding some control variables from a first stage that are included in a second stage, which is usually bad practice but can be done when there is strong theoretical justification for the exclusions. For instance, Carneiro, Heckman, and Vytlacil (2011) impose such a restriction on variables from the future affecting choices in the present.



integrated (coed) dorms and the StartUp dorm are in Table 4. We find no statistically significant treatment effects for either type of dorm on either ideation or intention, despite substantial F-statistics for our instruments between 50 and 300 depending on our specific (well above the heuristic threshold of 10 frequently recommended). This is strongly at odds with conclusions suggested by the raw OLS results without controls, which suggest strong positive associations between coed dorm assignments on ideas and intentions on men and negative associations for women. OLS with controls suggests large and statistically significant associations of 8 percentage points between StartUp dorm assignment and intention for men, at odds with the treatment effect estimate of -0.002 in our preferred specification. We note that the inclusion of extra controls has very little effect on IV point estimates, which is predictable given our identification strategy.

We next discuss effects of peer gender diversity, which are shown in Table 5. There is no variation in gender diversity at the roommate level, so we restrict our attention on gender diversity to gender distance between each individual and their next door neighbors. We do quite well at matching peer neighbor gender distance with out simulated instruments, with F-stats between 30 and 130 for our instruments depending on the specification. We find a large differential effect of -39 percentage points for women relative to men (significant at the 90% confidence level) of gender diversity on entrepreneurial intention, suggesting that gender integration significantly contributes to the gender gap in entrepreneurial intentions. The effect of peer gender diversity on intentions for men is statistically insignificant and positive. We see similar patterns with point estimates for the gender gap in entrepreneurial ideation, though these fall short of conventional standards of statistical significance. Though we are underpowered to say that gender integration encourages male entrepreneurship and discourages female entrepreneurship, we are powered to say that it has differentially negative effects on women compared to men on intentions.

Last, we discuss effects of academic peer diversity at the roommate level, which are shown in Table 6. We consider each academic trait of roommates separately, such that each coefficient is from a separate model. It is noteworthy that we do a worse job (as evidenced by F-statistics between 15 and 30) of predicting roommate academic traits than we do at predicting neighbor gender. This is likely because room genders are fixed ex ante before students are placed in them, such that our simulated instruments will closely match a student's neighbors' genders if we do well at simulating their room assignment. To match a student's roommates' academic traits,

Table 4: Effects of Dorm Assignments

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	IV	IV	IV
Panel 1: Effects on Ideas						
StartUp Dorm	0.015 (0.015)	0.032 (0.025)	0.045 (0.025)	0.017 (0.025)	0.029 (0.044)	0.034 (0.044)
Female x StartUp Dorm		-0.035 (0.031)	-0.046 (0.031)		-0.024 (0.052)	-0.034 (0.052)
Coed Dorm	0.015 (0.011)	0.090 (0.014)	-0.019 (0.019)	-0.040 (0.040)	-0.020 (0.077)	-0.029 (0.076)
Female x Coed Dorm		-0.133 (0.014)	0.038 (0.023)		-0.028 (0.088)	-0.022 (0.086)
1st Stage F-stat				299.247	50.241	50.524
Mean of DV	0.232	0.232	0.232	0.232	0.232	0.232
Panel 2: Effects on Intentions						
StartUp Dorm	0.046 (0.016)	0.062 (0.026)	0.082 (0.026)	0.006 (0.028)	-0.009 (0.047)	-0.002 (0.046)
Female x StartUp Dorm		-0.035 (0.033)	-0.046 (0.032)		0.020 (0.056)	0.012 (0.055)
Coed Dorm	0.032 (0.012)	0.144 (0.015)	0.003 (0.019)	-0.043 (0.042)	0.009 (0.080)	-0.004 (0.077)
Female x Coed Dorm		-0.199 (0.015)	0.011 (0.023)		-0.078 (0.092)	-0.079 (0.089)
1st Stage F-stat				299.247	50.241	50.524
Mean of DV	0.275	0.275	0.275	0.275	0.275	0.275
Extra controls	No	No	Yes	No	No	Yes
Observations	6919	6919	6919	6919	6919	6919

*Notes:* Linear probability models for effects of dorm assignments. All IV specifications control for expected values of simulated instruments. Robust standard errors in parentheses. Kleibergen-Paap F-stats reported for IV specifications. Specifications with extra controls contain all controls listed in Table 1 as well as intended major fixed effects.

Table 5: Effects of Peer Gender Diversity

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	IV	IV	IV
Panel 1: Effects on Ideas						
Neighbor Gender Distance	0.010 (0.014)	0.009 (0.023)	0.022 (0.023)	-0.128 (0.106)	0.006 (0.172)	0.010 (0.169)
Female x Neighbor Gender Distance		0.000 (0.029)	-0.005 (0.029)		-0.248 (0.216)	-0.244 (0.212)
1st Stage F-stat				130.161	34.519	34.614
Mean of DV	0.232	0.232	0.232	0.232	0.232	0.232
Panel 2: Effects on Intentions						
Neighbor Gender Distance	0.055 (0.015)	0.067 (0.025)	0.084 (0.024)	-0.025 (0.111)	0.193 (0.182)	0.197 (0.177)
Female x Neighbor Gender Distance		-0.022 (0.030)	-0.035 (0.030)		-0.406 (0.228)	-0.390 (0.221)
1st Stage F-stat				130.161	34.519	34.614
Mean of DV	0.275	0.275	0.275	0.275	0.275	0.275
Extra controls	No	No	Yes	No	No	Yes
Observations	6919	6919	6919	6919	6919	6919

*Notes:* Linear probability models for effects of peer gender distance of next door neighbors. All IV specifications control for expected values of simulated instruments. Robust standard errors in parentheses. Kleibergen-Paap F-stats reported for IV specifications. Specifications with extra controls contain all controls listed in Table 1 as well as intended major fixed effects.

we need to do well at simulating the student’s room *and* their roommates’ rooms, effectively requiring us to “guess right” for room assignments for two students instead of one. Despite our relatively weak instruments, we identify statistically significant (90th or 95th percentile depending on the inclusion of controls) effects of roommate verbal ACT distance on ideation of -3 percentage points. We find statistically insignificant positive effects of both ACT score distances on intentions, and of math score distances on ideation, with insignificant negative effects on both outcomes from roommate STEM distance. While many of our effects are imprecisely estimated, we note that roommate verbal distance not only has the most statistically significant effects, but it also likely has the most policy relevance as ACT verbal scores vary substantially more (8.9 SD) in our sample than ACT math scores (3.7 SD) or STEM intended major (0.50), as shown in Table 1.

Due to a lack of statistical power, we do not estimate effects of peers at further distances than roommates, for academic diversity, and next door neighbors, for gender diversity. Our exclusion of peer traits at further distances implicitly constrains their effects on outcomes to zero. If they do affect outcomes, they will bias the effects of nearer peers in the opposite

Table 6: Effects of Peer Academic Diversity

	(1)	(2)	(3)	(4)
	OLS	OLS	IV	IV
Panel 1: Effects on Ideas				
Roommate STEM Distance	0.011 (0.011)	0.009 (0.011)	-0.024 (0.272)	-0.081 (0.268)
1st Stage F-stat			17.393	17.194
Roommate ACT Math Distance	0.003 (0.002)	0.002 (0.002)	0.022 (0.034)	0.025 (0.033)
1st Stage F-stat			26.845	26.748
Roommate ACT Verbal Distance	-0.000 (0.001)	-0.001 (0.001)	-0.032 (0.016)	-0.030 (0.016)
1st Stage F-stat			16.015	16.165
Mean of DV	0.228	0.228	0.229	0.229
Panel 2: Effects on Intentions				
Roommate STEM Distance	0.014 (0.011)	0.012 (0.011)	-0.272 (0.294)	-0.314 (0.281)
1st Stage F-stat			17.393	17.194
Roommate ACT Math Distance	0.005 (0.002)	0.002 (0.002)	0.032 (0.035)	0.032 (0.034)
1st Stage F-stat			26.845	26.748
Roommate ACT Verbal Distance	0.001 (0.001)	-0.001 (0.001)	0.015 (0.016)	0.015 (0.016)
1st Stage F-stat			16.015	16.165
Mean of DV	0.273	0.273	0.272	0.272
Extra controls	No	Yes	No	Yes
Observations	6627	6627	6533	6533

*Notes:* Linear probability models for effects of peer academic distances of next door neighbors. All IV specifications control for expected values of simulated instruments. Robust standard errors in parentheses. Kleibergen-Paap F-stats reported for IV specifications. Specifications with extra controls contain all controls listed in Table 1 as well as intended major fixed effects. Individuals with no roommates are dropped for all specifications, and those with no simulated roommates are dropped from IV specifications.

direction of their effects because the traits of far peers are negatively correlated (mechanically) with the traits of near peers.<sup>18</sup> This mechanical negative correlation is small and is unlikely to produce large bias even if effects of distant peers are comparable in magnitude to effects of nearby peers, which is itself unlikely. Importantly, even if these effect estimates are biased in a strict sense, our results still give policy-relevant total effects of peer assignments that capture both the direct effect of the peer trait at a given distance on outcomes and the indirect effect of the peer trait at a given distance on outcomes via reducing (on average) the level of the trait among peers at further distances. We leave estimation of treatment effects of further peers to future work that is able to leverage more random variation in distant peers than we are able to with our identification strategy in our institutional context.

## 5. Conclusion

We find evidence that peer diversity, specifically with respect to neighbor gender and roommate verbal skills, is associated with *lower* entrepreneurial proclivity among students in a large public research university. We do not find statistically significant effects of STEM major diversity or peer math skill diversity. At the dorm level, we find no evidence for effects of either dorm entrepreneurial culture or gender integration, though gender-integrated dorm point estimates are consistent with the statistically significant effects of neighbor gender diversity on entrepreneurial intentions.

Our results are somewhat at odds with some existing literature that suggests exposure to course diversity increases entrepreneurship (Lazear, 2005) and diversity in teams enhances creativity and innovation (Wang, Kim, and Lee, 2016; Bolli, Renold, and Wörter, 2018; Schubert and Tavassoli, 2020). Our study is particularly strong in terms of identification of treatment effects because of the explicit randomization used in room assignments, but we have very limited qualitative information on the mechanisms whereby residential peers affect one another in our setting. It is possible that too much diversity reduces or changes the nature of student interactions, reducing information sharing and social encouragement that is conducive to entrepreneurship. It may also be that entrepreneurs actually benefit more from specialization than from diversity (Junkunc and Eckhardt, 2009; Khoshimov, Eckhardt, and Goldfarb, 2019). Qualitative research that investigates the nature of effective and ineffective peer interactions

---

<sup>18</sup>This is easy to imagine in a very small dorm. If there is a single STEM individual and they are next door to an individual, it must be the case that the individual's roommate is not STEM.

would shed light on this.

Our study contributes to the discussion of optimal room assignments in universities (and other settings such as office spaces). Existing work has found that assigning poor performing students to others as roommates can negatively affect their academic success (Sacerdote, 2001), especially for medium performers (Zimmerman, 2003). Other work suggests that exposure to racial diversity in roommates increases subsequent interracial interactions (Carrell, Hoekstra, and West, 2019). Our work considers entrepreneurship as an outcome, and suggests a role for clustering of students on academic skills and gender. Because the peer assignments that we find to be important for entrepreneurship do not overlap with those relevant for academic outcomes or hereditary inclusion, we consider our results to complement these other considerations of administrators in the question of optimal university dorm room assignments, which undoubtedly include an intention to improve all of these domains. A particularly relevant example of existing room assignment policies is that of learning communities. Many learning communities on the University of Wisconsin-Madison’s campus (and others) cluster students based on their academic interests (such as StartUp at UW-Madison) while others cluster students on gender and academic interests, such as WISE (Women in Science and Engineering). Our results suggest that such learning communities could play a substantial role in increasing campus entrepreneurship and reducing entrepreneurial gender gaps.

## References

- ABDULKADIROĞLU, A., J. D. ANGRIST, Y. NARITA, AND P. A. PATHAK (2017): “Research design meets market design: Using centralized assignment for impact evaluation,” *Econometrica*, 85(5), 1373–1432.
- ANGRIST, J., G. GRAY-LOBE, C. M. IDOUX, AND P. A. PATHAK (2022): “Still Worth the Trip? School Busing Effects in Boston and New York,” Discussion paper, National Bureau of Economic Research.
- ANGRIST, J., P. HULL, P. A. PATHAK, AND C. WALTERS (2021): “Credible school value-added with undersubscribed school lotteries,” *The Review of Economics and Statistics*, pp. 1–46.
- ANGRIST, J. D., G. W. IMBENS, AND D. B. RUBIN (1996): “Identification of causal effects using instrumental variables,” *Journal of the American statistical Association*, 91(434), 444–455.
- BELLONI, A., D. CHEN, V. CHERNOZHUKOV, AND C. HANSEN (2012): “Sparse models and methods for optimal instruments with an application to eminent domain,” *Econometrica*, 80(6), 2369–2429.
- BELLONI, A., V. CHERNOZHUKOV, AND C. HANSEN (2014): “High-dimensional methods and inference on structural and treatment effects,” *Journal of Economic Perspectives*, 28(2), 29–50.
- BOLLI, T., U. RENOLD, AND M. WÖRTER (2018): “Vertical educational diversity and innovation performance,” *Economics of Innovation and New Technology*, 27(2), 107–131.
- BORUSYAK, K., AND P. HULL (2020): “Non-random exposure to exogenous shocks: Theory and applications,” Discussion paper, National Bureau of Economic Research.
- BROUWER, J., A. FLACHE, E. JANSEN, A. HOFMAN, AND C. STEGLICH (2018): “Emergent achievement segregation in freshmen learning community networks,” *Higher Education*, 76(3), 483–500.
- CARNEIRO, P., J. J. HECKMAN, AND E. J. VYTLACIL (2011): “Estimating marginal returns to education,” *American Economic Review*, 101(6), 2754–81.

- CARRELL, S. E., M. HOEKSTRA, AND J. E. WEST (2019): “The impact of college diversity on behavior toward minorities,” *American Economic Journal: Economic Policy*, 11(4), 159–82.
- DE CHAISEMARTIN, C. (2017): “Tolerating defiance? Local average treatment effects without monotonicity,” *Quantitative Economics*, 8(2), 367–396.
- GNEEZY, U., AND A. RUSTICHINI (2000): “Pay enough or don’t pay at all,” *The Quarterly Journal of Economics*, 115(3), 791–810.
- GOTTFRIED, M. A., AND A. HARVEN (2015): “The effect of having classmates with emotional and behavioral disorders and the protective nature of peer gender,” *The Journal of Educational Research*, 108(1), 45–61.
- GREENE, P. G., M. M. HART, E. J. GATEWOOD, C. G. BRUSH, AND N. M. CARTER (2003): “Women entrepreneurs: Moving front and center: An overview of research and theory,” *Coleman White Paper Series*, 3(1), 1–47.
- GUZMAN, J., AND A. O. KACPERCZYK (2019): “Gender gap in entrepreneurship,” *Research Policy*, 48(7), 1666–1680.
- HARRIS, C. (2022): “Interpreting Instrumental Variable Estimands with Unobserved Treatment Heterogeneity: The Effects of College Education,” *Working Paper*.
- HAUSMAN, N. (2012): “University innovation, local economic growth, and entrepreneurship,” *US Census Bureau Center for Economic Studies Paper No. CES-WP-12-10*.
- HIRANO, K., AND G. W. IMBENS (2004): “The propensity score with continuous treatments,” *Applied Bayesian modeling and causal inference from incomplete-data perspectives*, 226164, 73–84.
- ICE (2016): “Stem designated degree program list,” <https://www.ice.gov/sites/default/files/documents/document/2016/stem-list.pdf>, [online; accessed 2021-07-23].
- IMBENS, G. W. (2000): “The role of the propensity score in estimating dose-response functions,” *Biometrika*, 87(3), 706–710.
- IMBENS, G. W., AND J. D. ANGRIST (1994): “Identification and estimation of local average treatment effects,” *Econometrica*, 62(2), 467–475.



- JUNKUNC, M. T., AND J. T. ECKHARDT (2009): “Technical specialized knowledge and secondary shares in initial public offerings,” *Management Science*, 55(10), 1670–1687.
- KHOSHIMOV, B., J. T. ECKHARDT, AND B. GOLDFARB (2019): “Abandonment of the applicants signal: Grades and entrepreneurship,” in *Academy of management proceedings*, vol. 2019, p. 12909. Academy of Management Briarcliff Manor, NY 10510.
- LAZEAR, E. (2005): “Entrepreneurship,” *Journal of Labor Economics*, 23(4), 649–680.
- OOSTERBEEK, H., AND R. VAN EWIJK (2014): “Gender peer effects in university: Evidence from a randomized experiment,” *Economics of Education Review*, 38, 51–63.
- ROESSNER, D., J. BOND, S. OKUBO, AND M. PLANTING (2013): “The economic impact of licensed commercialized inventions originating in university research,” *Research Policy*, 42(1), 23–34.
- ROSENBAUM, P. R., AND D. B. RUBIN (1983): “The central role of the propensity score in observational studies for causal effects,” *Biometrika*, 70(1), 41–55.
- ROTHAERMEL, F. T., S. D. AGUNG, AND L. JIANG (2007): “University entrepreneurship: a taxonomy of the literature,” *Industrial and corporate change*, 16(4), 691–791.
- RUBIN, D. B. (1978): “Bayesian inference for causal effects: The role of randomization,” *The Annals of statistics*, pp. 34–58.
- (1980): “Randomization analysis of experimental data: The Fisher randomization test comment,” *Journal of the American statistical association*, 75(371), 591–593.
- (1990): “Comment: Neyman (1923) and causal inference in experiments and observational studies,” *Statistical Science*, 5(4), 472–480.
- SACERDOTE, B. (2001): “Peer effects with random assignment: Results for Dartmouth roommates,” *The Quarterly journal of economics*, 116(2), 681–704.
- SCHUBERT, T., AND S. TAVASSOLI (2020): “Product innovation and educational diversity in top and middle management teams,” *Academy of Management Journal*, 63(1), 272–294.
- SØRENSEN, J. B., AND A. J. SHARKEY (2014): “Entrepreneurship as a mobility process,” *American Sociological Review*, 79(2), 328–349.

- WANG, X.-H. F., T.-Y. KIM, AND D.-R. LEE (2016): “Cognitive diversity and team creativity: Effects of team intrinsic motivation and transformational leadership,” *Journal of business research*, 69(9), 3231–3239.
- WRIGHT, M., AND P. MUSTAR (2019): *Student start-ups: The new landscape of academic entrepreneurship*, vol. 1. World Scientific.
- WRIGHT, M., D. S. SIEGEL, AND P. MUSTAR (2017): “An emerging ecosystem for student start-ups,” *The Journal of Technology Transfer*, 42(4), 909–922.
- ZHAO, C.-M., AND G. D. KUH (2004): “Adding value: Learning communities and student engagement,” *Research in higher education*, 45(2), 115–138.
- ZIMMERMAN, D. J. (2003): “Peer effects in academic outcomes: Evidence from a natural experiment,” *Review of Economics and statistics*, 85(1), 9–23.
- ZÖLITZ, U., AND J. FELD (2021): “The effect of peer gender on major choice in business school,” *Management Science*, 67(11), 6963–6979.

## Appendix A: Tiebreaker Randomization Check

The key to our research design is the randomization of room assignment tiebreakers, which are used in conjunction with student preferences to assign students to rooms. If tiebreakers are not actually random, our identification strategy is not valid. We note that Housing staff has told us that tiebreakers are random, while also telling us that their official policy is that staff judgment in assignments trumps any deference to the random allocation mechanism. It therefore seems to us that Housing has no incentive to doctor the random numbers to benefit some groups over others, even if they did have an interest in giving certain groups particularly advantageous room assignments.

To address the possibility of nonrandom tiebreaker numbers, we perform randomization checks by regressing random tiebreakers on the observed student characteristics described in Table 1. We report  $\beta$  coefficients from year-specific regressions of the form

$$r_i = X_i\beta + \epsilon_i,$$

where we test the null hypothesis that  $\beta = 0$  for every element of  $X_i$  other than the constant.

Results are shown in Table A.1. Unsurprisingly, we find no evidence of randomization failures; there is one coefficient that is significant at the 95% confidence level and 5 that are significant at the 90% confidence level. Considering there are 48 coefficients presented, the number of statistically significant associations is consistent with randomization of tiebreakers. We are further encouraged that the statistically significant effects differ across years, which would not occur if Housing were systematically biased toward or against certain student populations.

Table A.1: Tiebreaker Randomization Check

	2016 (1)	2017 (2)	2018 (3)	2019 (4)
Female	-5.905 (131.0)	-51.91 (143.4)	37.73 (147.4)	-23.36 (168.0)
STEM	-41.78 (132.3)	11.59 (143.2)	136.7 (148.9)	-283.5* (164.1)
ACT Math	13.78 (20.62)	4.638 (23.26)	-35.98 (22.35)	19.41 (24.75)
ACT Verbal	-0.368 (8.728)	3.323 (7.935)	8.424 (9.968)	-0.507 (10.65)
Nontraditional	-168.2 (503.6)	-302.0 (477.9)	272.1 (789.3)	579.4 (757.2)
In-State	134.9 (138.4)	141.1 (153.6)	-39.96 (157.7)	330.3* (182.0)
First Generation	41.40 (168.2)	16.55 (188.3)	152.5 (176.3)	-102.1 (220.3)
Asian	-91.90 (332.9)	448.0 (390.7)	616.6* (373.4)	-291.4 (398.2)
Black	963.6** (476.1)	-547.0 (561.6)	259.2 (668.4)	764.1 (956.3)
Hispanic	32.78 (407.4)	255.8 (422.1)	-36.78 (442.4)	-110.2 (505.9)
White	266.7 (241.6)	328.6 (286.6)	548.8* (286.8)	0.155 (317.8)
International	413.3 (319.3)	103.6 (373.5)	186.1 (363.4)	385.6 (400.4)
Constant	4246.4*** (710.7)	4272.6*** (784.1)	4762.7*** (794.8)	4162.1*** (888.1)
Observations	2116	1860	1571	1277

*Notes:* Predictive associations between baseline student characteristics and tiebreakers, by year.

\*, \*\*, and \*\*\* denote significance at the 90, 95, and 99% confidence levels.

## Appendix B: Room Assignment Monotonicity

As discussed in Section 3.3., a threat to identification is monotonicity violations for our simulated instrumental variables. Differences between simulated treatment assignments and actual treatment assignments arise either because some students are not assigned according to the mechanism we identify as best-fitting the data, and because some students are assigned in this way for their summer assignments, but request room changes for reasons such as conflict with roommates. The first group, students our assignment mechanism makes mistakes on, are likely individuals who either have unobserved room priorities in the rooms they receive (always takers or never takers), or they are random mistakes for whom the compliers-defiers assumption is likely particularly valid.

The second group of students who dislike their room(mate) assignment and request a new room are more problematic. It is possible to imagine a particularly cantankerous type of student who, upon being assigned a roommate of any type, takes issue with their roommate and requests a move. If such individuals have abnormal treatment effects from assignment to peers (for instance, if they are antisocial and have opposite signed effects from the general population for all peer traits), this will contribute to bias in treatment effect estimates if they are able to move to opposite-type peers when they request a room switch.

We have two arguments for monotonicity violations from cantankerous students being unlikely to invalidate our estimates. First, per interviews with Housing staff, students who request room moves are not allowed to request any particular alternative room(mate), they are placed near their old room in an open spot if one is available. It follows that they are not systematically likely to receive the opposite peer exposure from their initial assignment, except for the small mechanic effect from their prior room being unable to be their new room. It follows that even cantankerous students who reject initial assignments are often not defiers with respect to their peer or dorm treatment assignments.

Even for the subset of these students who are defiers, the compliers-defiers version of the monotonicity assumption requires only that there be sufficiently many compliers who share treatment effects with defiers to cancel them out. This means we need individuals who relate to their peers similarly to the way individuals who request room changes do in the complier population. The initial assignment takeup rate is informative about the number of defiers there may be in our population, as we are primarily concerned about students rejecting their initial

assignment and receiving a different realized assignment (we are not concerned about students having an initial assignment that doesn't match our simulated assignments). Statistics on room take-up are shown in Table B.1. From this table, it looks like room assignment rejections are extremely rare at just over 2% of assignments. If every single individual who rejects their initial assignment has abnormal effects from exposure to peers, it is sufficient for us for the most similar individuals among the great mass of compliers to have overlapping treatment effects. Given that there appear to be at least 50 compliers for every defier, we feel confident that this overlap condition is satisfied.

Table B.1: Initial Room Assignment Take-up by Student Type

	Room Takeup Rate (1)
Total	0.978
Female	0.977
STEM	0.980
Nontraditional	0.962
In-State	0.982
First Generation	0.972
Asian	0.976
Black	0.951
Hispanic	0.987
White	0.979
Other Race	0.980
International	0.983
No Roommate	0.955
Random Room Assignment	0.980
Random Roommate Assignment	0.966
Observations	6919

*Notes:* Percentage of students of each type whose actual room is the same as their summer assignment.