# Peer Effects in University Housing:
# Evidence from Fuzzy Central Assignment[*]

Chao Fu[†]    Jesse Gregory[‡]

Clint Harris[§]    Victoria Prowse[¶]

This version: October 26, 2022
First version:

## Abstract

We estimate effects of roommates, neighbors, and dorms on academic outcomes at a large public four-year university. To address selection, we instrument for realized room assignments with simulated offers generated by a room assignment mechanism, while controlling for expected offers implied by the same mechanism. Our candidate assignment mechanisms fail to perfectly replicate the assignment offers made by the central planner, so we select the mechanism that best replicates assignments via a data-driven model selection procedure. We find that assignment to gender-integrated dorms increases 4-year graduation by 10 percentage points for men, with no significant effects on women, while finding consistent (statistically insignificant) results for exposure to female neighbors within dorms. At the roommate level, we find that STEM roommates have significantly smaller negative effects on four-year graduation for other STEM students than they do for non-STEM students. Our findings suggest that universities with 70% gender-integrated housing (such as the university we study) could increase their male 4-year graduation rates by 3 percentage points by implementing full gender-integration in all dorms, with no offsetting negative effects on women.

**Keywords**: Peer Effects, College Majors, Central Assignment, Model Selection, Simulated Instruments

**JEL Classification**: C52, D47, I20

---

[*]Institutional Review Board approvals were obtained from The University of Wisconsin (2021-0422) and Purdue University (IRB-2021-1024).

[†]Department of Economics, University of Wisconsin-Madison; cfu@ssc.wisc.edu.

[‡]Department of Economics, University of Wisconsin-Madison; jmgregory@ssc.wisc.edu.

[§]Corresponding Author: Wisconsin School of Business, University of Wisconsin-Madison; clint.harris@wisc.edu.

[¶]Department of Economics, Purdue University; vprowse@purdue.edu.

# 1 Introduction

A large number of young people attend universities, many of whom live in university housing. Their assignments to dorms and peers are of policy interest because complementarities between individuals and their dorms or peers allows for the possibility of alternative living arrangements that may improve outcomes on average, or even for all groups. Because many aspects of these living arrangements are determined by university administrators according to a randomized procedure, a large literature has developed to consider which types of peers positively affect other peers, with a particular focus on nonlinear effects which enable aggregate gains from counterfactual peer assignments. In this paper, we estimate effects of roommate, neighbor, and dorm characteristics on a range of academic outcomes in the setting of a public four-year university by leveraging intent-to-treat simulated peer assignments generated by a data-driven room assignment mechanism.

A large literature investigates effects of peers in university housing on academic and vocational outcomes. However, this literature overwhelmingly leverages data from private universities (Sacerdote, 2001; Stinebrickner and Stinebrickner, 2006; Marmaros and Sacerdote, 2006; Sacerdote, 2011) and military academies (Carrell et al., 2013, 2019; Jones and Kofoed, 2020) where the randomization of room assignments is well-understood, but which might not be representative of peer effects in more common settings.[1] We provide evidence on effects of roommates, neighbors, and dorms in student housing at a large public four year institution in the United States, merging administrative data on pre-existing student characteristics, room assignments, and outcomes with internal-use data from University Housing on dorm and roommates preferences. Specifically, we follow 24,265 incoming freshmen at the University of Wisconsin-Madison from 2016-2019 who submit housing preferences the summer before their first year on campus.

It is possible that the limited attention to large public universities in the literature is due to room assignment complexities that invalidate commonly-applied empirical strategies that condition on observed room assignment strata indicators as in the research described above. We find that room assignment procedures in our public four year institutional setting are substantially more complex and less consistently documented than those described in the literature.[2] We overcome this challenge by applying and extending recent methodological innovations by Abdulkadiroğlu et al. (2017), Borusyak and Hull (2020), and others that instrument for treatment assignments with treatment offers, while controlling for expected values of offers. These methods have been used to great effect to estimate effects, for instance, of school value added (Angrist et al., 2017), charter schools (Abdulkadiroğlu et al., 2017), and travel distance to school (Angrist et al., 2022). We are the first (to our knowledge) to apply these methods to estimation of peer effects in a university setting.

Unlike recent applications in the school choice literature of these methods, we are not able to replicate the observed room assignments with a parsimonious room assignment mechanism, eliminating our ability to control for expected values of treatment offers. Our key methodological insight is that the results of Abdulkadiroğlu et al. (2017) and Borusyak and Hull (2020) permit treatment effect estimation of simulated treatments when controlling for expected simulated treatments *regardless of whether simulated treatment offers correspond to realized treatment offers*. It

---

[1] A well-known exception is Foster (2006), who extensively discusses serious selection problems with respect to roommates in her institutional setting.

[2] For instance, in interviews with housing staff, we were informed that university housing staff exercises judgment while doing room assignments, intentionally attempting to improve roommate match quality according to unspecified rules. One example given was that they avoid assigning students from the same high school as roommates if the high school is small. There is neither a requirement that staff always perform this correction nor a standard metric for determining whether a high school is small.

follows that the dorm or peer assignments implied by any arbitrary room assignment mechanism are valid (but perhaps irrelevant) instruments for realized treatment assignments when controlling for students' expected dorm or peer assignments implied by the same mechanism. In other words, we construct "intent-to-intent-to-treat instruments", where we as researchers intend for our assignment mechanism to produce treatment assignments that match those intended by the central assigner. Realized assignments may differ from our simulated assignments either because of lack of take-up from the central planner when they make assignment offers that conflict with those that we simulate, or because of lack of take-up from students when they refuse assignment offers made by the central planner.

While any room assignment mechanism is likely to produce conditionally independent instruments, there is no requirement that it will produce relevant instruments that reliably predict realized dorm or peer assignments. Formally, this constitutes a setting with infinitely many potential instruments and a finite number of observations. To construct relevant instruments, we simulate room assignments according to a small number of candidate mechanisms that satisfy parsimony constraints, and we test the accuracy of each mechanism against the assignment offers (not realized assignments). To avoid complicating inference for treatment effects via pretesting, we select the room assignment mechanism that best rationalizes the randomly-generated (and therefore independent of relevant peer traits) lottery room assignment lottery tiebreakers of students' peers implied by assignment offers, rather than attempting to match dorm or peer characteristic variables of realized peers directly. Model selection methods such as those we use are potentially useful to school (etc.) choice applications with researcher ignorance regarding true assignment mechanisms as well as in the more general framework of Borusyak and Hull (2020) where there is uncertainty regarding the economic model generating the data.

We estimate effects of roommate ACT scores, roommate STEM status, neighbor gender, and dorm-level gender integration. Our results on dorms are particularly well-powered, and are made possible by our school-choice style identification style that explicitly leverages a substantial amount of between-dorm identifying variation. We find that assignment to coeducational dorms (those with no partitions between genders) increases four-year graduation rates by over 10 percentage points for men (90% CI) while having no significant effect on women. Given that just under 70% of students in our sample live in coed dorms, linear extrapolationn suggests that our university could potentially increase four-year graduation for male students by 3 percentage points, with no negative offsets to women, by converting all dorms to be fully coeducational. This is consistent with previous findings that women and girls in educational settings positively effect college graduation rates (Hill, 2017), high school graduation rates and test scores (Lavy and Schlosser, 2011), and primary school reading and math scores (Gottfried and Harven, 2015).

In addition to effects at the dorm level, we estimate effects of neighbor gender by leveraging more granular variation in exposure to opposite-gender peers made possible by the existence of coeducational dorms. To do this, we manually geocode dormrooms in three dimensions using university blueprints. This allows us to construct a matrix of (Manhattan) distances in inches between each room and every other room on the same floor. We define neighbors for individuals in each room as those individuals that inhabit the nearest three rooms, allowing us to investigate effects of peers at a more granular level than past work, such as Foster (2006) who found null results of nearby peers when considering coarser wing-level variation in peers. We that exposure to female neighbors increases the probability of graduating with a STEM major by 19 percentage points, with suggestive (but insignificant) evidence that this effect is larger for men. We also find statistically insignificant evidence that exposure to female peers positively effects freshmen retention and four-year graduation for men with no effects on women — which we view as complementary to our similar estimates of effects of coeducational dorms.

We also consider effects of roommates. We estimate effects of peer STEM interest (based on intended majors), peer verbal ACT scores, peer math ACT scores, and distances between individuals values of these traits and those of their roommates. We find that assignment to a STEM roommate has a 7 percentage point more positive effect for STEM individuals relative to non-STEM individuals on four-year graduation rates, with opposite effects of similar magnitude on students graduating with a STEM major. We also find that a standard deviation (8.8 point) increase in ACT verbal score distance narrowly falls short (p-value 0.13) of statistical significance in negatively affecting freshmen retention by 7 percentage points. These findings broadly suggest potential benefits of clustering students together based on academic abilities and interests, rather than exposing them to academically diverse peers.

The paper proceeds as follows: Section 2 describes the empirical strategy; Section 3 describes the data; Section 4 presents the results; and Section 5 concludes.

## 2  Empirical Strategy

### 2.1  Room Choice

Raw comparisons between individuals with different peers or dorms will not identify treatment effects of peer or dorm traits of interest if individuals choose dorms or rooms based on their unobserved academic ability or other relevant traits. For instance, if students who excel academically tend to inhabit particular dorms, we will observe positive associations between measures of peer academic ability and academic outcomes such as grades and graduation rates that do not reflect peer effects. To address selection bias, our empirical strategy leverages the randomness in the room assignment mechanism used by University Housing to calculate and control for expected treatments, which is sufficient to address omitted variable bias from otherwise unmodeled determinants of assignments (Rosenbaum and Rubin, 1983; Hirano and Imbens, 2004; Borusyak and Hull, 2020). To facilitate the discussion, we adapt notation from Abdulkadiroğlu et al. (2017) (AANP) to describe room allocations. A room choice problem assigns individuals $i = 1, 2, ..., N$ in set $\mathcal{I}$ to room-spots (beds) indexed by $s$, with $s = 1, 2, \ldots, S$. There are $N$ total students and $S$ total room spots, where a particular feature of our setting is that $N = S$ because the assignment of interest is a room-spot rather than a dorm.

A room assignment mechanism $\varphi$ incorporates student preferences over rooms, room priorities over students, and a randomly assigned tiebreaker to place students into room-spots. Each student $i$ in has a preference ordering over rooms, $\succ_i^\varphi$, where $A \succ_i^\varphi B$ means that the mechanism $\varphi$ treats student $i$ as though they prefer dorm $A$ to dorm $B$. Meanwhile, each room $s$ gives mechanism-specific priority $\rho_{is}^\varphi$ to student $i$ with $\rho_{is}^\varphi < \rho_{ij}^\varphi$ indicating that mechanism $\varphi$ treats the room as preferring student $i$ over $j$. Each student has a type denoting nonrandom determinants of their room assignments indicated by $\theta_i^\varphi = (\succ_i^\varphi, \boldsymbol{\rho}_i^\varphi)$, where $\boldsymbol{\rho}_i^\varphi$ is a vector that contains $i$'s priority in each room. Finally, each student is assigned a randomly-generated tiebreaker number, $r_i$ in $\{1, 2, ..., N\}$, which denotes their position in the order, and this tiebreaker is assumed to be observed by the researcher. A student's room-spot assignment as determined by mechanism $\varphi$ is given by $\mu_i^\varphi(\theta^\varphi, r) \in \{1, 2, \ldots, S\}$, with $i$'s realized assignment $\mu_i = \mu_i^*(\theta^*, r)$ being defined as the assignment determined by the true (potentially unknown to the researcher) allocation mechanism, denoted by $\varphi = *$.

A contribution of this paper is to extend the setup of AANP to the case where we allow for arbitrary mechanism-specific determinants of assignments ($\theta^\varphi$) for any arbitrary mechanism that maps from these determinants to assignments, $\mu^\varphi(\theta^\varphi, r)$. As described by AANP, room assignments can be simulated for any mechanism that satisfies Equal Treatment of Equals (ETE), which requires

that a mechanism gives all individuals of the same type the same assignment probabilities for all rooms. Our framing presents ETE as an information problem for the researcher rather than a fairness problem for the central planner; all mechanisms with assignments given by $\mu_i^\varphi(\theta^\varphi, r)$ satisfy equal treatment of equals if $\mu_i^\varphi(\cdot)$ is known and $(\theta^\varphi, r)$ are observed. Because room assignments can be simulated for any mechanism defined by a researcher that only uses inputs observed by the researcher, it follows that the effects of treatments implied by such assignments are also identified by controlling for expected treatments.[3]

The room choice problem we describe assigns students to rooms, so we could, in principle, estimate room value added by comparing individuals assigned to a particular room on campus to others (treating an arbitrarily chosen room as the omitted treatment of reference). However, very few (usually two) students are assigned to each room at a given time, so we would be woefully underpowered to identify treatment effects for rooms. More importantly, any room-level treatment effects will lack policy-relevance outside of our immediate institutional context. Finally, we expect ex ante that rooms themselves are of limited importance, with the substantive mechanisms through which room assignments affect outcomes being characteristics of peers (roommates and neighbors) and the dorm in which the room resides. We therefore describe individual $i$'s treatment status $D_{ik}(\mu^\varphi(\theta^\varphi, r))$ for treatment $k = 1, \ldots, K$ as defined by the room allocations of all students under mechanism $\varphi$.

## 2.2 Treatment Effect Estimation

Our empirical strategy leverages conditional randomization in the room assignment procedure. A general representation of the models we will estimate is

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

where $Y_i$ denotes an outcome of interest such as 4 year graduation or having a STEM major, $D_i^\varphi = D_{ik}(\mu^\varphi(\theta, r))$ is the value of treatment $k$ (with $k$ subscripts omitted for clarity) such as the share of $i$'s roommates under a given mechanism with different major STEM status, $\mathbb{E}[D_i^\varphi|\theta^\varphi]$ is the expected value of this simulated 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; Hirano and Imbens, 2004; 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 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. Specifically, our empirical strategy is a stratified randomized research design with selection on observables.[4]

---

[3]It may well be that an arbitrary assignment mechanism imagined by a researcher will have no connection to reality. In this case, we would expect to have identified treatment effects of zero on all outcomes of all simulated dorm traits and simulated peer traits.

[4]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.

There are three noteworthy approaches for estimating peer effects that will not reliably provide consistent estimates across institutional settings. 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 a sufficient statistic for $\theta_i^*$. We are concerned that this will not be the case in our setting, specifically that students placed into dorms early in the assignment procedure may rank their dorm highly and be placed near to one another within the dorm - introducing nonrandom dorm-preference clustering within dorms.[5] The next approach is to control for every possible permutation of dorm preferences, assuming that preferences are sufficient statistics for $\theta_i^*$. There are 21 dorms in our application 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 ($\theta_i^*$ also contains room priorities over students, which may be functions of other observed or unobserved student characteristics). The final approach is to calculate $\mathbb{E}[D_i^*|\theta^*]$ 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 et al. (2017). In our application, we do not know the mechanism used by Housing, so we cannot calculate $\mathbb{E}[D_i^*|\theta^*]$.

While we cannot calculate $\mathbb{E}[D_i^*|\theta^*]$, we can calculate $\mathbb{E}[D_i^\varphi|\theta^\varphi]$ for an arbitrary mechanism that we specify. 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 lacks immediate policy relevance, despite its causal interpretation, because it is the effect of a dorm or peer trait as implied by a mechanism specified by the researcher, not the effect of an actual dorm or peer trait. However, following the same arguments as we give above for the identification of the effect of $D_i^\varphi$ on $Y_i$, we can also identify the effect of $D_i^\varphi$ on $D_i$ from

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

where $\gamma_D^\varphi$ is the effect of assignments as implied by mechanism $\varphi$ on actual assignments $D_i$. 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 instrumental variables assumptions, which we will discuss presently.

## 2.3   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,

---

[5]Our concern is validated in our application by the relationship between peer priority number and own priority number 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.

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 2.4.

The instrumental variables exclusion restriction requires that peer or dorm traits implied by room assignment offers as determined by an assignment mechanism only affect outcomes via the corresponding realized peer or dorm traits. Two straightforward examples of possible violations come to mind. One is that individuals other than the student will obtain information on their offers (regardless of takeup) and treat the student differently, for instance if the university assigns course instructors or time-slots jointly with residences, for instance to offset disappointment students may feel from not receiving preferred housing assignments. We discussed such data uses with adminstrative staff, and they emphasized that the residence hall data is not used for any other purposes. The other natural concern is that room assignment offers may induce students to interact with intended roommates during the summer (for instance to coordinate on living arrangements) and that these summer interactions may have effects on outcomes regardless of whether a student takes up their room(mate) assignment. We assume that such effects negligible, and we add that because over 97% of students accept room offers (Table B.1), any such effects are approximately equivalent to being effects of realized peers.% [6]

A less straightforward potential violation of the exclusion restriction involves violation of the Stable Unit Treatment Value Assumption (SUTVA), which requires that treatment assignment means the same thing for all individuals (Rubin, 1978, 1980, 1990).If SUTVA is violated, a student could receive a particular type of peer (e.g. a Math major roommate) for the true realization of the random tiebreaker vector, a different type of peer (e.g. a Biology major roommate) for a counterfactual realization of the random tiebreaker vector, and have the same treatment status as modeled (e.g. a STEM major roommate) with different outcomes — constituting a violation of the exclusion restriction. A special case of this problem can occur due to peer spillovers, for instance if having a STEM roommate who has STEM neighbors is a substantively different treatment than having a STEM roommate who has non-STEM neighbors, and these are not modeled as distinct treatment statuses. We expect our analyses to be rife with both of these sorts of SUTVA violations.

A related issue involving unmodeled heterogeneity is that of heterogeneous treatment effects, which is typically addressed by the monotonicity assumption of Imbens and Angrist (1994). Monotonicity requires that simulated peer or dorm assignments induce individuals into corresponding realized treatment statuses without ever inducing any individuals out of those statuses. In our context, monotonicity implies that zero students have opposite-signed mismatch between their simulated peers and actual peers for all values of their simulated peer traits. This effectively rules out the possibility of random mistakes in our simulation of room assignments.[7] It also rules out the possibility that some students are discontented with their initial room (or roommate) assignment regardless of what it is, with the result being that they always request a new assignment before the start of the year. We expect such students to be rare, but they nonetheless constitute a violation of monotonicity.

We doubt that SUTVA or monotonicity are satisfied in our setting, so we make the alternative

---

[6]If other university settings have similar room offer takeup, then the technical exclusion violation here is a policy-relevant one, where our IV estimates will provide (ceteris paribus) valid predictions of peer effects in other settings.

[7]For instance, if a random 10% of students have actual peer assignments that differ from their simulated assignments (either because we make mistakes in matching the true assignment algorithm, or because administrative personel make mistakes in implementing it), we would expect 1% (10% of 10%) of students to be defiers for binary peer traits, because their simulated peers would differ from realized peers in all states of the world.

6

(closely related) assumption of net uniform unordered monotonicity (Harris, 2022). This assumption extends the multiple treatment unordered monotonicity condition of Heckman and Pinto (2018) to the case of unobserved treatments (allowing for SUTVA violations), while also incorporating the compliers-defiers condition of De Chaisemartin (2017) to the case of multiple treatments (allowing for monotonicity violations). The "uniform unordered monotonicity" part of this condition imposes that being assigned simulated peers with a modeled composite trait (e.g. STEM major) makes realized assignment to peers with all unmodeled versions of the trait (Biology major, Math major, etc) weakly more likely for every student. The "net" part of this condition allows for there to be defiers who respond to simulated treatment assignments by being less likely to have some or all versions of the treatment as long as there exist compliers who share their treatment effects, effectively cancelling out the contribution of defiers to local average treatment effects estimates. In simple terms, we assume that our simulated peer assignments will either match realized peer assignments with respect to both modeled and unmodeled peer traits, or that they will fail to match realized peer assignments *nonsystematically*. Insofar as defiers arise because of mistakes we make in replicating the true room assignment mechanism, we expect this assumption to be particularly valid.[8] Insofar as some defiers may be students who reject their initial assignment for behavioral reasons, we assume that these students are rare enough to have overlapping treatment effects with a (possibly very abnormal) equal mass subset of the complier population. We present evidence in Appendix B that such defiers are likely quite rare, accounting for at most 2% of the student population, which gives us confidence that compliers exist to cancel them out.

## 2.4 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 et al. (2012) and Belloni et al. (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

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

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

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

---

[8]Individuals whose room assignments are invariant to our simulation (e.g. individuals who are nonrandomly assigned to particularly advantageous rooms for unobserved reasons) will be always takers or never takers, not defiers.

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.[9] 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 from proposed mechanisms 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 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

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

$D_i^{**}$ 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.[10]

As an example of case (3b) above where a deterministic mechanism produces weak instruments, consider the deterministic-on-observables mechanism $\varphi = \theta$ that satisfies $\mu = \mu^*(\theta, r) = \mu^\theta(\theta^\theta)$, where sufficient individual characteristics are contained in $\theta^\theta$ 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

---

[9]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.

[10]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.

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 $\mu_i^\theta(\theta^\theta, r) = \mu_i^\theta(\theta^\theta, r')$ for all $i$ for any alternative set of tiebreakers $r'$. In this case, the first stage in (2) collapses to

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

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 $\mathbb{E}[D_i^\theta | \theta^\theta]$ 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.[11] Second, we require that all candidate mechanisms 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 mechanisms that that are monotonic functions of random tiebreakers. Specifically, we require that the preference rank of the dorm a student is assigned to is monotonically increasing in the value of their random tiebreaker, ceteris paribus (early ranked students get their preferences over late students). Similarly, we also require that the position in mechanism-determined order of the room a student is assigned is monotonicially increasing in the value of their random tiebreaker conditional on dorm assignments (early ranked students are in "earlier" rooms, where earlier is defined by each mechanism, for instance, by a lower room number). 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. .

## 3 Data and Institutional Details

### 3.1 Data Description

To estimate peer effects, we use data on undergraduate students who live in dorms from 2016 to 2019 at the University of Wisconsin-Madison, a large, public research university, with outcomes covering the same timespan. This administrative data contains information 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, test scores, and intended majors. Additionally, we have a wealth of outcome variables, including 4-year graduation rates (for the 2016 cohort), freshmen retention (for all but the 2019 cohort), second year and graduation majors, and course grades.Importantly, our data contains the residence hall preference lists provided by students, which are 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. There are 24,265 undergraduate students who live in dorms in their first fall which are assigned during the preceding

---

[11]This restriction was incentive compatible, as conducting this interview was much easier than coding a room assignment mechanism.

summer, of whom 17,972 have second year outcomes (retention and second year majors for those retained) and 5,902 have four year outcomes (graduation and graduation majors for those who graduate). 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 submitting their dorm application late) because they are not included in the university's random room allocation process that is central to our empirical strategy.

The peer traits that we calculate are neighbor share female, roommate share STEM major distance, average roommate ACT math score, average roommate ACT verbal score, and distances between individuals own statuses and those of their roommates for STEM status, ACT math, and ACT verbal. For each of these, distances are calculated by taking the absolute value difference between an individual and the peer trait of each relevant peer, then averaging over these distances.[12] 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 Math is defined by the score on the corresponding test, while ACT Verbal is the sum of the scores on the ACT Reading test and the ACT English test.[13] ACT scores come from administrative data, and we impute ACT scores for students with missing scores using official SAT to ACT conversion tables for students with SAT scores, and we impute scores using AP test count, AP test score average, high school rank, high school class size, and high school GPA for students with neither ACT scores nor SAT scores.

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, the inhabitants of which are treated as neighbors for the student(s) in the room.

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). As we show in Table 1, about 70% of dorms at Unviersity of Wisconsin-Madison are fully gender-integrated, suggesting substantial scope for policy improvements if coeducational dorms are found to have either positive or negative effects on student outcomes.

Our empirical strategy will control 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 intended-major fixed effects (2-digit Classification of Instructional Programs codes). Summary statistics for the variables we use are available in Table 1.

---

[12]In words, we calculate the average distance between individual $i$ and their peers, not the distance between individual $i$ and their average peer.

[13]The SAT to ACT conversion table makes use of the same coding of the ACT Verbal score.

Table 1: Descriptive Statistics

|                              | Mean | SD |
|                              | (1) | (2) |
|------------------------------|--------|-------|
| Freshmen Retention           | 0.957  | 0.204 |
| 4 Year Graduation            | 0.715  | 0.452 |
| Second Year Major STEM       | 0.350  | 0.477 |
| 4 Year STEM Grad             | 0.411  | 0.492 |
| Coed Dorm                    | 0.697  | 0.460 |
| Female                       | 0.531  | 0.499 |
| STEM                         | 0.435  | 0.496 |
| ACT Math                     | 28.634 | 3.741 |
| ACT Verbal                   | 57.749 | 8.944 |
| Nontraditional Student       | 0.016  | 0.125 |
| In-State Student             | 0.556  | 0.497 |
| First Generation Student     | 0.183  | 0.387 |
| Asian                        | 0.071  | 0.257 |
| Black                        | 0.018  | 0.132 |
| Hispanic                     | 0.056  | 0.230 |
| White                        | 0.707  | 0.455 |
| Other Race                   | 0.149  | 0.356 |
| International                | 0.091  | 0.288 |
| Fall 2016 Cohort             | 0.243  | 0.429 |
| Fall 2017 Cohort             | 0.245  | 0.430 |
| Fall 2018 Cohort             | 0.253  | 0.435 |
| Fall 2019 Cohort             | 0.259  | 0.438 |
| No Roommate                  | 0.039  | 0.193 |
| Random Room Assignment       | 0.691  | 0.462 |
| Random Roommate Assignment   | 0.387  | 0.487 |
| Observations                 | 24265  |       |

*Notes:* Means and standard deviations for outcomes, peer traits, and controls for all first year students in dorms..

## 3.2 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 peers assigned to them may be assigned according to the random mechanism. Similarly, the university allows incoming 1st-year students to designate a preferred roommate with whom they will share a room. These students will not be randomly assigned to roommates, but they will still be randomly assigned to dorms unlike those assigned via steps 1-4. Generally, these details suggest that we will have substantially more random variation for identifying dorm effects and neighbor effects than for roommate effects, as roommates are more rarely assigned randomly. 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). With the caveat that Housing staff judgment dominates other considerations, Housing staff conveyed to us the following information regarding the dorm assignment mechanism (Step 6):

1. Assign each student on campus a random number.

2. Assign each set of predetermined roommates the minimum value of their random numbers.[14]

3. Order students in ascending order of random numbers.

4. Proceed through the list from top to bottom, assigning each student to their favorite dorm if possible, otherwise their next favorite, etc, with each student's assignment being resolved before considering subsequent students' preferences.

---

[14]This systematically advantages students with predetermined roommates with respect to dorm assignments.

5. Assign each student to rooms in an unspecified way.

The assignment mechanism described by Housing staff for dorms corresponds to the Random Serial Dictator (RSD) assignment mechanism for students who are randomly assigned to rooms.[15]

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 use the methods described in Section 2.4 to adjudicate between plausible room assignment mechanisms we describe in Section 4.1, with results on mechanism performance shown in Table 2.

# 4 Results

We present model selection results in Section 4.1. We present treatment effect estimates of peers and dorms in Section 4.2.

## 4.1 Mechanism Selection

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 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 $r_i = 1, 2, ..., N$ without reference to other individuals. Next, we consider the Boston mechanism, which assigns individuals in order $r_i = 1, 2, ..., R$ 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 first assigns students to dorms via Boston *if and only if they are placed in an unpopular dorms*, otherwise skipping them. After all unpopular dorms are filled, this mechanism starts over and places remaining students in dorms

---

[15]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 and these students are treated as though they have room-specific preferences for their assigned room.

according to RSD.[16]

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 (for instance displayed next to 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, with ascending z (floor) order.[17]

Standard tests of instrument strength would estimate equations such as (3) and consider the F-statistic associated with 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 $j = 1, 2, ..., J$ for each dorm assignment mechanism,

$$
\begin{aligned}
Dorm_{i,1} &= \alpha^\varphi_{Dorm} Dorm^\varphi_{i,1} + u^\varphi_{i,1} \\
Dorm_{i,2} &= \alpha^\varphi_{Dorm} Dorm^\varphi_{i,2} + u^\varphi_{i,2} \\
&\vdots \\
Dorm_{i,J} &= \alpha^\varphi_{Dorm} Dorm^\varphi_{i,J} + u^\varphi_{i,J},
\end{aligned}
\tag{7}
$$

where we emphasize that the constant is constrained to 0 and all equations in the model are constrained to share a single slope coefficient. $\alpha^\varphi_{Dorm}$ has the attractive property of giving the weighted average over of all dorms of 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

$$
\begin{aligned}
r_{i,0} &= \alpha^\varphi_r r^\varphi_{i,0} + \delta^\varphi_{i,0} + u^\varphi_{i,0} \\
r_{i,1-3} &= \alpha^\varphi_r r^\varphi_{i,1-3} + \delta^\varphi_{i,1-3} + u^\varphi_{i,1-3} \\
r_{i,4-7} &= \alpha^\varphi_r r^\varphi_{i,4-7} + \delta^\varphi_{i,4-7} + u^\varphi_{i,4-7},
\end{aligned}
\tag{8}
$$

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 $\delta^\varphi_i$ denoting simulated dorm fixed effects. We report dorm assignment mechanism accuracy, $\alpha^\varphi_{Dorm}$, in Panel 1 of Table 2 and room assignment mechanism accuracy, $\alpha^\varphi_r$, 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

---

[16]The gist of this mechanism is that it will place people in unpopular dorms if they rank them relatively highly, regardless of their other dorm preferences. We suggested this mechanism to Housing after failing to match assignments with both RSD and Boston, and they told us it loosely approximates the sort of ad hoc deviations they occassionally make from the assignments implied by RSD. As shown in Table 2, it performs quite poorly.

[17]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.

Table 2: Mechanism Selection

|  | Random 1 | Random 2 | 2016 | 2017-2019 | Total |
| --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) |
| Panel 1: Dorm Assignments |  |  |  |  |  |
| RSD | 0.865 | 0.878 | 0.961 | 0.852 | 0.872 |
|  | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Boston | 0.802 | 0.826 | 0.852 | 0.808 | 0.811 |
|  | (0.002) | (0.002) | (0.002) | (0.001) | (0.001) |
| Boston → RSD | 0.676 | 0.692 | 0.742 | 0.679 | 0.684 |
|  | (0.002) | (0.002) | (0.003) | (0.002) | (0.001) |
| Observations | 8427 | 8427 | 3968 | 12886 | 16854 |
| Panel 2: Peer Assignments |  |  |  |  |  |
| Room ID | 0.500 | 0.432 | 0.413 | 0.444 | 0.437 |
|  | (0.006) | (0.006) | (0.008) | (0.005) | (0.004) |
| Room # | 0.515 | 0.445 | 0.425 | 0.459 | 0.451 |
|  | (0.006) | (0.006) | (0.008) | (0.005) | (0.004) |
| Geographic | 0.511 | 0.437 | 0.426 | 0.450 | 0.445 |
|  | (0.006) | (0.006) | (0.008) | (0.005) | (0.004) |
| Observations | 7659 | 7672 | 3705 | 11626 | 15331 |

*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.

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 instrumental variables models described by equations (2) and (1), but we are encouraged by the cross-sample consistency in accuracy-maximizing mechanisms. We find that Random Serial Dictator is the best dorm assignment mechanism, with Room # order being the best room assignment mechanism — though all room assignment mechanisms have very similar accuracy. We note that we have smaller samples for accuracy measures because we only test accuracy for individuals who are randomly assigned to rooms, while the sample shrinks further for room accuracy checks because they rely on simulated and realized roommates, which are missing for individuals in single person rooms.

Our selected mechanisms produce simulated and expected dorm and peer treatment assignment statuses described in Table 3. There are two main takeaways from this table. First, our expected, simulated, and realized assignments have almost identical means for all variables.[18] Second, re-

---

[18]This is approximately true by construction as there is only one pool of students, though different realizations of assignments may differ in which students are placed in single-person rooms or rooms with more than two people, where their traits will receive less weight. Additionally, Housing sometimes changes the number of students in a room between summer assignments and fall assignments. Our realized treatments take room occupancy values values from fall assignments while our expected and simulated treatments take them from summer assignments. More extreme differences can occur with nonlinear peer traits, such as the peer trait distances measures (these could be set to zero for all students, in principle).

alized 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.[19]

Table 3: Treatment Assignment Description

|  | Mean (1) | SD (2) |
|---|---|---|
| Coed Dorm | 0.697 | 0.460 |
| Simulated Coed Dorm | 0.698 | 0.459 |
| Expected Coed Dorm | 0.698 | 0.370 |
| Neighbor Female Share | 0.521 | 0.368 |
| Simulated Neighbor Female Share | 0.518 | 0.371 |
| Expected Neighbor Female Share | 0.518 | 0.289 |
| Roommate STEM | 0.434 | 0.489 |
| Simulated Roommate STEM | 0.430 | 0.485 |
| Expected Roommate STEM | 0.430 | 0.382 |
| Roommate STEM Distance | 0.412 | 0.485 |
| Simulated Roommate STEM Distance | 0.421 | 0.484 |
| Expected Roommate STEM Distance | 0.423 | 0.381 |
| Roommate ACT Math | 28.649 | 3.677 |
| Simulated Roommate ACT Math | 28.648 | 3.648 |
| Expected Roommate ACT Math | 28.645 | 2.982 |
| Roommate ACT Math Distance | 3.643 | 2.858 |
| Simulated Roommate ACT Math Distance | 3.664 | 2.856 |
| Expected Roommate ACT Math Distance | 3.689 | 2.361 |
| Roommate ACT Verbal | 57.806 | 8.648 |
| Simulated Roommate ACT Verbal | 57.760 | 8.614 |
| Expected Roommate ACT Verbal | 57.759 | 6.877 |
| Roommate ACT Verbal Distance | 8.803 | 7.209 |
| Simulated Roommate ACT Verbal Distance | 8.853 | 7.172 |
| Expected Roommate ACT Verbal Distance | 8.920 | 5.958 |
| No Roommate | 0.039 | 0.193 |
| No Simulated Roommate | 0.047 | 0.211 |
| Observations | 24265 | |

*Notes:* Summary statistics for treatment variables as well as simulated treatments and expected simulated treatments from best-fitting mechanism (RSD assignments to dorms, Room # order assignments to rooms).

---

[19]A non-stratified RCT would produce approximately zero variance in expected treatments. In our application, expected treatments explain well over half of the variation in realized treatments, suggesting that most of the variation in peer and dorm assignments is driven by selection.

## 4.2 Treatment Effect Estimates

In this section we present treatment effect estimates for dorm and peer treatment assignments. In general, we estimate instrumental variables models that take the form

$$Y_i = \beta_{0k} + \beta_{Dk}D_{ik} + \beta_k \mathbb{E}[D_{ik}^\varphi|\theta^\varphi] + \epsilon_{ik},$$
$$D_{ik} = \gamma_{0k}^\varphi + \gamma_{Dk}^\varphi D_{ik}^\varphi + \gamma_k^\varphi \mathbb{E}[D_{0ik}^\varphi|\theta] + u_{ik}^\varphi.$$

by two stage least squares, giving estimates of linear marginal effects of peer/dorm assignment $k$ on the probability of outcome $Y$. We investigate effects of dorm coeducational status, neighbor gender, and roommate ACT scores and intended majors.

In general, we estimate models with a single treatment variable (or with an own trait interaction), rather than combining all treatments into a single model. If peer traits are correlated (they are), these specifications still give policy relevant total effects of a particular peer or dorm trait on a particular outcome.[20] Such estimates are sufficient statistics for predicting the effect of counterfactual increases/decreases in assignments to a particular type or dorm or peer for a given outcome. Richer models with multiple treatments do not immediately inform likely effects of policies that change dorm/peer assignments with respect to single (for instance statistically significant) trait, because they estimate partial effects of assignment to many types of dorms and peers *conditional on other traits being held fixed*. In other words, our specifications can be readily used to answer questions of the form "what would happen if we assigned $x$ more individuals to $y$ type dorms/peers?", while specifications with multiple treatments cannot (additional information about covariances between peer and dorm traits is needed to use them for such predictions).

The question stated in the preceding paragraph is not policy relevant if relevant counterfactual policies are not feasible. In general, inelastically supplied treatments with homogeneous effects cannot produce benefits by being redistributed, and the information needed to effectively redistribute them is not available if their effects are estimated in (misspecified) homogeneous effects models. Main effects of our treatments, therefore, are policy relevant to the situation in which dorms or peers with particular traits are not inelastically supplied. The implication of this for our application is that homogenous treatment effect specifications inform university policies that increase the availability of particular dorm or peer treatments, for instance by converting non-gender-integrated dorms into gender-integrated dorms, or by admitting more students with particular characteristics to the university (or assigning more of such students to university housing). In general, we consider such single effect specifications to be relevant for dorm effects, but to not be particularly relevant for peer effects.[21]

---

[20]Strictly speaking, these specifications are likely to produce exclusion restriction violations in the sense that the modeled trait is correlated with unmodeled traits that are contained in the error term. If the covariances between the modeled trait and unmodeled traits in our dataset are representative of those in the population, these total effect estimates retain external validity. This issue is ubiquitous in instrumental variables estimation. For instance, the effects of charter schools estimated by Abdulkadiroğlu et al. (2017) similarly are partially driven by in-sample covariances between charter school status and other unmodeled school characteristics (insofar as other characteristics have effects). One principled way to dispute findings from such studies (including ours) is to suggest that these in-sample covariances are significantly different in the research sample than they are in the population.

[21]Even if a given university were to admit additional students with a particular trait (e.g. high verbal ACT scores) based on our findings or other insights, it is likely that any net benefits accrued by the university would be offset by negative effects on these students' second-most-preferred universities. It is possible that universities competing over students who produce desirable peer effects may improve market efficiency (for instance by increasing college attendance rates among such students due to increases in college-provided financial aid), but we expect these effects to be second order relative to benefits from selection-on-gains types of policies that leverage heterogeneous treatment effects to assign peers to students they will particularly benefit.

We show in Table 4 the effects of living in coeducational (highly gender-integrated) dorms on a range of academic outcomes. Of particular note, we find positive effects for men of 10.7 percentage points from assignment to a coed dorm relative to a non-coed dorm. We find no evidence of positive or negative effects for women for four-year graduation, or for either men or women for other outcomes of interest. One policy implication from this result is that increasing the number of coeducational dorms on campus should be expected to increase male four-year graduation rates without substantially affecting women. An alternative implication is that similar gains could be made by assigning all men to the existing coeducational dorms, while housing women in the less gender-integrated dorms as needed. While it is possible that the effects of coed dorms are unrelated to their gender compositions (which would raise concerns about external validity) or that they are nonlinearly related to their gender compositions, we are doubtful that such a counterfactual arrangement would preserve the mechanisms that are present within our sample through which coed dorms increase graduation rates for men.[22]

For peer effects specifications, we estimate models that include peer traits alongside own-trait interactions with peer traits. In these models, the main effect is of relatively little interest, while the interaction answers questions of the form "how much would outcome $y$ change if we assigned a peer with trait $x$ to an individual with that same trait instead of an individual without it?". For binary peer traits, the coefficient on the uninteracted term in these specifications gives the increase in an outcome for the group that lacks the trait. If there is a particular interest in increasing the outcome for this particular group (for instance to address preexisting inequalities), then these effects are policy-relevant. If only aggregate changes in outcomes are policy interest, then a significant (economically or statistically) effect for a particular group produces no relevant implications unless some other group has a different effect (with this information conveyed by the estimated effect of the interaction term).

We show effects of the share of neighbors that are female in Table 5. We find point estimates that corroborate our findings for coeducational dorms, suggesting that exposure to female neighbors improves freshmen retention and four-year graduation for men, with no offsetting negative effects for women (though these estimates are all statistically insignificant at conventional levels). We also find that female neighbors increase STEM majors among individuals who graduate, though these effects are not systematically different for men and women. The lack of statistical significance on differences in effects between men and women suggests limited scope to increase aggregate STEM graduation rates, but the statistically significant main effect does suggest that increases in gender segregation may reduce the STEM major gender gap (increasing STEM majors among graduating women and decreasing them among graduating men). We note that we have significantly less power to estimate effects of neighbors relative to dorms, as our preferred mechanism is approximately 87% accurate for dorm assignments, while being only 45% accurate for peer assignments (shown in Table 2). This is reflected in the dramatically lower F-statistics for neighbor traits as compared to dorm assignments.

We next turn to effects of academic traits of roommates, with results shown in Table 6. We find that assignment to a STEM roommate statistically insignificantly reduces four-year graduation rates for non-STEM individuals, but that it has statistically-significantly smaller negative effects (by 7 percentage points) on STEM individuals. A statistically significant interaction effect in the opposite direction of a statistically insignificant main effect may seem to be of questionable policy relevance, but we reiterate our point above that the interaction effects are substantially more policy relevant than main effects because they are sufficient to predict aggregate effects of counterfactual

---

[22]The extreme case in which "coed" dorms are exclusively inhabited by men strikes us as a stark (and humorous) example of a policy that would fail the Lucas (Lucas, 1976) critique.

peer assignments. Specifically, we would expect that assigning all STEM students to each other as peers would produce aggregate four-year graduation rates 3.1 percentage points higher than the alternative extreme case of assigning no STEM students to each other as peers, and 1.5 percentage points higher than the existing allocation, in which 52% of students with STEM intended majors have roommates with STEM intended majors. We also find that STEM peers have significantly smaller effects on the probability of graduating with a STEM major on STEM individuals than they do for non-STEM individuals. A plausible explanation for this finding is that because incoming freshmen with STEM intended majors have relatively high propensities to graduate with STEM majors, they are relatively inelastic to any relevant treatment, including peer assignments. In other words, this effect can be explained by a constant effect of STEM peers on the latent utility associated with graduating with a STEM major that manifests differently in terms of STEM major probabilities because of different latent utility distributions for different groups. Nonetheless, our findings still suggest that assigning STEM individuals to each other as roommates would reduce aggregate STEM graduation majors by 3.7 percentage points relative to the alternative extreme case of assigning no STEM students to each other as peers while reducing aggregate STEM graduation majors by 1.8 percentage points relative to the existing allocation.

We have substantially less power to estimate effects of roommate traits, both because of our room assignment mechanism's poor accuracy for peer assignments relative to dorm assignments, and because only 39% of students in our sample are randomly assigned to roommates, approximately 10% of whom (4% of the total sample) receive no roommate due to being placed in a single-occupancy room. This is readily seen in our relatively low F-statistics in these models, which range from 32 for outcomes for which we observe a large portion of the sample to below 2 for outcomes that we only observe for the 2016 incoming freshmen cohort. Our point estimates on ACT scores are in line with some past results such as those of Zimmerman (2003) which find relatively larger effects for verbal scores relative to math scores, though in our application these effects are not significant at conventional levels.

Interactions between own traits and peer traits allow for one form of policy-relevant heterogeneity in treatment effects. We also consider the distances between individuals' academic traits and those of their roommates. This allows us to directly answer questions regarding whether students should be clustered among peers of their own type or whether they should be exposed to peers who are different from them. For binary traits (such as gender and STEM major), an individual's peer trait distance is a deterministic function of the peer trait of interest and the interaction between their own trait and the peer trait — implying that estimates of effects of peer trait distances will be similar to those obtained in the models with interactions described above. For traits where an individual's peers may have either higher or lower values of a trait than an individual, specifications that model effects of distances between traits convey unique information about potentially policy relevant nonlinearities in effects.

To see the potential value of individual-peer trait distance specifications relative to specifications with interactions, it is helpful to consider a DGP where the distance between an individual's trait and those of her peers drives outcomes for all types of individuals, but there are no linear effects for any group. To fix ideas for this example, we proceed with the simplifying assumption that each individual's expected peer trait is the same as their own trait, and we consider individuals at only two values of the trait (though their peers may have any value). Example data from this sort of model for an outcome $Y$ and peer trait $x_j$ is shown in Panel 1 of Figure 1. With such a DGP, it is tempting to conclude that the estimated effect of the own trait peer trait interaction in a model that includes this term alongside a peer trait main effect will yield a positive coefficient, as high trait individuals are particularly strongly affected by being assigned high trait peers (these matches have low trait distances). This ignores, however, that the identification strategy relies

on controlling for expected peer traits. Controlling for expected peer traits produces a coefficient on the own-trait-peer-trait interaction that is driven by the comparison of outcomes between low trait individuals who are assigned to peers with higher or lower *than expected* traits and high trait individuals who are assigned to peers with higher or lower *than expected* traits. In other words, even though high trait individuals' peers have high traits more often than low trait individuals' peers, they do not have higher-than-expected traits more often. Controlling for expected peers in this scenario conditions out the variation in peer traits that would allow a negative peer trait distance effect to be captured by an own-trait-peer-trait interaction.[23]

Table 4: Dorm Effects

|  | Freshmen Retention | | 4 Year Graduation | | 4 Year STEM Grad | |
|---|---|---|---|---|---|---|
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| Coed Dorm | 0.001 | 0.014 | 0.041 | 0.107* | 0.038 | 0.097 |
|  | (0.010) | (0.017) | (0.034) | (0.063) | (0.032) | (0.059) |
| Female × Coed Dorm |  | -0.022 |  | -0.104 |  | -0.088 |
|  |  | (0.021) |  | (0.074) |  | (0.070) |
| 1st Stage F-stat | 1544.509 | 302.335 | 923.755 | 142.790 | 766.666 | 107.780 |
| Mean of outcome | 0.957 | 0.957 | 0.715 | 0.715 | 0.411 | 0.411 |
| Observations | 17972 | 17972 | 5902 | 5902 | 4627 | 4627 |

*Notes:* Effects of dorm assignments estimated via 2SLS instrumenting for realized assignments with simulated assignments. All specifications control for expected values of simulated instruments, controls listed in Table 1, and intended major fixed effects. Robust standard errors in parentheses. */**/*** denote significance at the 90, 95, and 99 percentage confidence levels.

Table 5: Effects of Peer Gender

|  | Freshmen Retention | | 4 Year Graduation | | 4 Year STEM Grad | |
|---|---|---|---|---|---|---|
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| % Neighbor Female | 0.030 | 0.054 | 0.087 | 0.161 | 0.194** | 0.248 |
|  | (0.028) | (0.039) | (0.087) | (0.135) | (0.095) | (0.162) |
| Female × % Neighbor Female |  | -0.052 |  | -0.131 |  | -0.105 |
|  |  | (0.056) |  | (0.175) |  | (0.198) |
| 1st Stage F-stat | 326.074 | 73.800 | 199.220 | 55.524 | 132.878 | 24.885 |
| Mean of outcome | 0.957 | 0.957 | 0.715 | 0.715 | 0.411 | 0.411 |
| Observations | 17972 | 17972 | 5902 | 5902 | 4627 | 4627 |

*Notes:* Effects of gender composition of next door neighbors estimated via 2SLS instrumenting for realized peers with simulated peers. All specifications control for expected values of simulated instruments, controls listed in Table 1, and intended major fixed effects. Robust standard errors in parentheses. */**/*** denote significance at the 90, 95, and 99 percentage confidence levels.

The results on distances between individual's traits and their peer traits are shown in Table 7. We do not find statistically significant impacts for any trait considered. Interestingly, we find substantially larger effects of roommate ACT Verbal score distance than we do for main effects of ACT verbal scores in Table 6. However, the estimated effect of roommate ACT Verbal distance falls narrowly short of conventional confidence levels, with a p-value of 0.13, with no significant

---

[23]The intuition that an interaction term will capture effects of distances between individuals' traits and those of their peers is correct for binary treatments for which all individuals have the same expected peer traits — an ex ante unlikely scenario that is at odds with the positive standard deviations on expected peer traits shown in Table 1.

Table 6: Effects of Peer Academic Traits

| | Freshmen Retention | | 4 Year Graduation | | 4 Year STEM Grad | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Roommate STEM | -0.139 | -0.142 | -0.128 | -0.169 | 0.289 | 0.331 |
| | (0.093) | (0.093) | (0.247) | (0.246) | (0.298) | (0.294) |
| $\times$ STEM | | 0.005 | | 0.071** | | -0.086** |
| | | (0.010) | | (0.036) | | (0.038) |
| 1st Stage F-stat | 32.054 | 16.102 | 23.629 | 11.874 | 15.694 | 7.958 |
| Roommate ACT Math | -0.014 | 0.013 | -0.003 | 0.024 | -0.008 | -0.029 |
| | (0.032) | (0.042) | (0.105) | (0.123) | (0.124) | (0.144) |
| $\times$ ACT Math | | -0.001 | | -0.001 | | 0.001 |
| | | (0.001) | | (0.002) | | (0.002) |
| 1st Stage F-stat | 59.053 | 29.198 | 27.239 | 13.656 | 16.648 | 8.367 |
| Roommate ACT Verbal | 0.036 | 0.032 | 0.302 | 0.432 | -0.413 | -0.463 |
| | (0.039) | (0.047) | (0.223) | (0.293) | (0.456) | (0.597) |
| $\times$ ACT Verbal | | 0.000 | | -0.002 | | 0.001 |
| | | (0.000) | | (0.002) | | (0.002) |
| 1st Stage F-stat | 32.430 | 16.243 | 8.318 | 4.085 | 1.867 | 0.863 |
| Mean of outcome | 0.957 | 0.957 | 0.718 | 0.718 | 0.405 | 0.405 |
| Observations | 16986 | 16986 | 5602 | 5602 | 4351 | 4351 |

*Notes:* Effects of peer academic traits of roommates estimated via 2SLS instrumenting for realized peers with simulated peers. All specifications control for expected values of simulated instruments, controls listed in Table 1, and intended major fixed effects. STEM is a binary indicator for having a STEM intended major, while ACT scores are normalized to have unit standard deviations. Robust standard errors in parentheses. Individuals with no roommates or no simulated roommates are dropped. */**/*** denote significance at the 90, 95, and 99 percentage confidence levels.

Figure 1: Example Data for Peer Trait Distances

Panel A                                             Panel B
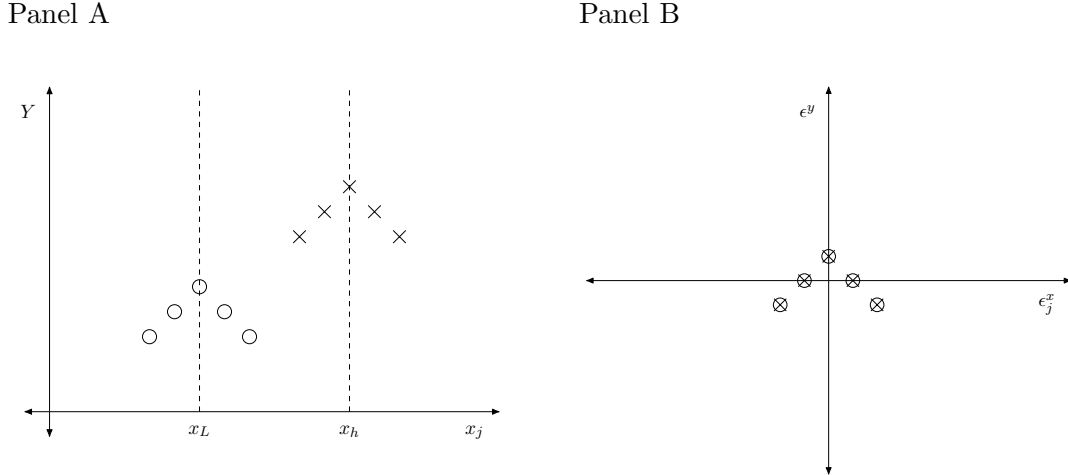


*Notes:* Graphical representation of example data where outcomes are determined only by the individual's own trait and distances between an individual's trait and the same trait of their peers. Panel A shows a scatter plot of raw data for two types of individuals, with type $L$ defined as $x_i = x_L$ shown as circles and type $H$ defined as $x_i = x_H$ shown as x's, under the simplifying assumption that $\mathbb{E}[x_j | x_i] = x_i$ (each point could in principle represent any number of observations, with the caveat that the expected value for each group be unchanged). Panel B shows the Frisch-Waugh-Lovell residuals after controlling for the expected value of peer traits. It is apparent from Panel B that a linear regression of the outcome on peer traits, controlling for expected peer traits, will fail to uncover an economically significant relationship (regardless of interactions included in the model), while a distance-between-traits specification will succeed.

difference in effects estimated in a model that also includes a linear term for average roommate ACT verbal scores.[24] Taking the narrowly insignificant effect of standard deviation increases in peer ACT verbal distance on freshmen retention at face value produces a prediction that assigning all students to roommates with identical verbal ACT scores (an approximately feasible policy) would increase freshmen retention by 8.8 percentage points, which would (implausibly) predict aggregate freshmen retention of 104.5%. The potential for large effects from feasible policy interventions regarding this trait is compelling, and we feel warrants future research on the role of peer verbal skill differences in determining outcomes. A potential explanation for such effects is that students communicate more easily with peers with similar verbal skills, and that effective communication between peers fosters incresased educational success and increases the amenity value of staying in school.

Table 7: Effects of Peer Academic Diversity

|  | Freshmen Retention | | 4 Year Graduation | | 4 Year STEM Grad | |
| --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| Roommate ACT Math Distance | 0.009 | 0.006 | -0.016 | -0.016 | 0.076 | 0.069 |
|  | (0.038) | (0.042) | (0.137) | (0.138) | (0.130) | (0.133) |
| Roommate ACT Math |  | -0.013 |  | -0.004 |  | 0.000 |
|  |  | (0.035) |  | (0.105) |  | (0.128) |
| 1st Stage F-stat | 56.669 | 12.865 | 22.753 | 8.233 | 23.163 | 8.141 |
| Roommate ACT Verbal Distance | -0.072 | -0.065 | -0.057 | -0.122 | -0.098 | -0.186 |
|  | (0.048) | (0.049) | (0.241) | (0.299) | (0.264) | (0.349) |
| Roommate ACT Verbal |  | 0.023 |  | 0.305 |  | -0.427 |
|  |  | (0.040) |  | (0.232) |  | (0.439) |
| 1st Stage F-stat | 27.730 | 10.120 | 7.835 | 1.513 | 3.974 | 0.811 |
| Mean of outcome | 0.957 | 0.957 | 0.718 | 0.718 | 0.405 | 0.405 |
| Observations | 16986 | 16986 | 5602 | 5602 | 4351 | 4351 |

*Notes:* Effects of ACT score distances between roommates estimated via 2SLS instrumenting for realized peers with simulated peers. All specifications control for expected values of simulated instruments, controls listed in Table 1, and intended major fixed effects. ACT scores are normalized to have unit standard deviations, so distance effects give predictions of a peer with a one standard deviation higher or lower ACT score. Robust standard errors in parentheses. Individuals with no roommates or no simulated roommates are dropped. */**/*** denote significance at the 90, 95, and 99 percentage confidence levels.

---

[24]The minimal impact of also including main effects of peer traits alongside distances on estimates of distance effects is predictable. For students whose expected peer's trait is the same as their own trait, the correlation between peer trait distances and peer traits is mechanically driven toward zero when conditioning on expected traits — distant peers cannot have traits that are systematically higher than expected or lower than expected.

# 5    Conclusion

This paper reports estimates of peer and dorm effects using a sample drawn from University of Wisconsin-Madison, a large public four-year university. We overcome substantial empirical challenges resulting from selection into dorms and peers by implementing an intent-to-intent-to-treat instrumental variables strategy that, vitally, is robust to misspecification of the unviersity's room assignment mechanism. Our methodological contribution is to extend methods commonly employed in the school choice literature to a setting where researcher-proposed treatment assignment mechanisms fail to replicate realized assignments. We perform a straightofrward model selection exercise which identifies a best-performing mechanism without requiring the computationally demanding simulation of expected treatments under alternative mechanisms. This method has potential to be of use in similar school (etc) choice settings in which researchers are unaware of assignment mechanisms used by central planners (or in which central planners do not explicitly use a consistent mechansim). More generally, our model selection procedure can be used to choose between alternative economic models which produce different predictions for realized and expected treatments in implementations of the instrumental variables methods described by Borusyak and Hull (2020).

Substantively, we find that highly gender-integrated dorms increase four-year graduation rates for men, while STEM peers reduce four-year graduation rates less for STEM individuals than for non-STEM individuals. We find promising suggestive evidence (falling short of conventional significance thresholds) that female neighbors increase freshmen retention for men and that high distance between ACT verbal scores reduces freshmen retention. Taken together, our findings suggest significant potential increases in freshmen retention and four-year graduation from approximately costless counterfactual room assignment mechanisms. Specifically, we predict that converting all dorms to be highly-gender integrated and exclusively assigning STEM students to each other as roommates would increase aggregate four-year graduation rates by approximately 3 percentage points relative to baseline.[25]

Our findings on gender effects complement a large literature that generally finds positive effects of girls and women on the educational outcomes of their peers. We acknowledge that our findings on the effects of coeducational dorms are identified off of the approximately ten highly-coeducational dorms contained within our sample. With this small number of dorms, it is possible that an alternative amenity shared by coed dorms on campus is the true driver of outcomes. Future research estimating effects of gender-integration in dorms on outcomes at other universities would address this concern.

Our finding of aggregate benefits from clustering STEM individuals together as roommates complements some similar findings in the peer effects literature. For instance, Booij et al. (2017) find that group homogeneity of tutorial (study) groups improved performance among students at an economics and business school. Similarly, Duflo et al. (2011) find that students in Kenya benefitted from being in classroom environments with similar-ability peers. We broadly replicate these findings for effects of homogeneity of academic interests (STEM vs. non-STEM roommates) in addition to academic ability (distances between roommates ACT verbal scores), while also showing that they hold in residential settings in addition to classrooms and study groups. We expect these results to be particularly externally valid in other settings that use similar measures of STEM academic interests (and in which student's majors within STEM are similar) and verbal skills, such as other universities in the United States. We see no reason to expect substantial differences in effects for alternative measures of STEM interests or verbal skills, but studies at other universities could

---

[25]Driven by a 30 percentage point increase in coed dorm assignments affecting 50% of the student population (males) by at a rate of 0.1, and a 48 percentage point increase in STEM roommate propensities affecting 43% of the student population (STEM individuals) at a rate of 0.07.

shed light on this. Along these lines, if students were to intentionally alter their verbal test scores or (more plausibly) their reported intended majors to the university in an attempt to receive a particular peer assignment, this could invalidate the predictions from estimates. If universities do not publically (or even privately, as in the case of the university we study) reveal the details of their room assignment mechanisms, we expect that this sort of gaming by students will be unlikely.

We are particularly optimistic regarding the external validity of our findings. First, we estimate effects in a large public four-year university. Even if our findings are not generalizable to educational residential accommodations in other contexts (which they may be), we expect them to be generalizable to similar universities, which educate and house a large share of undergraduate students. Second, many of the counterfactual room assignment policies we discuss occur naturally within the support of our data. As discussed by Booij et al. (2017) and starkly investigated by Carrell et al. (2013), extrapolation outside of the support of data relies on potentially invalid functional form assumptions. Our estimates suggest potentially large effects from increases in the number of highly gender-integrated dorms as well as from clustering STEM students with each other as roommates. Even larger effects may be possible with clustering on verbal ACT scores, though are estimates are insufficient to predict effects of this sort of policy with certainty. Corroboration of these point estimates in other settings would potentially reduce this uncertainty, with substantial social benefits, though the consistency between our results and those of Duflo et al. (2011) and Booij et al. (2017) may be sufficient to motivate trials of such policies by university housing adminsitrators.

# References

**Abdulkadiroğlu, A.**, **Angrist, J.D.**, **Narita, Y.**, and **Pathak, P.A.** (2017). Research design meets market design: Using centralized assignment for impact evaluation. *Econometrica*, 85(5): 1373–1432

**Angrist, J.**, **Gray-Lobe, G.**, **Idoux, C.M.**, and **Pathak, P.A.** (2022). Still worth the trip? school busing effects in boston and new york. Working Paper 30308, National Bureau of Economic Research. doi:10.3386/w30308

**Angrist, J.D.**, **Hull, P.D.**, **Pathak, P.A.**, and **Walters, C.R.** (2017). Leveraging lotteries for school value-added: Testing and estimation. *The Quarterly Journal of Economics*, 132(2): 871–919

**Belloni, A.**, **Chen, D.**, **Chernozhukov, V.**, and **Hansen, C.** (2012). Sparse models and methods for optimal instruments with an application to eminent domain. *Econometrica*, 80(6): 2369–2429

**Belloni, A.**, **Chernozhukov, V.**, and **Hansen, C.** (2014). High-dimensional methods and inference on structural and treatment effects. *Journal of Economic Perspectives*, 28(2): 29–50

**Booij, A.S.**, **Leuven, E.**, and **Oosterbeek, H.** (2017). Ability peer effects in university: Evidence from a randomized experiment. *The review of economic studies*, 84(2): 547–578

**Borusyak, K.** and **Hull, P.** (2020). Non-random exposure to exogenous shocks: Theory and applications. Technical report, National Bureau of Economic Research

**Carrell, S.E.**, **Hoekstra, M.**, and **West, J.E.** (2019). The impact of college diversity on behavior toward minorities. *American Economic Journal: Economic Policy*, 11(4): 159–82

**Carrell, S.E.**, **Sacerdote, B.I.**, and **West, J.E.** (2013). From natural variation to optimal policy? the importance of endogenous peer group formation. *Econometrica*, 81(3): 855–882

**De Chaisemartin, C.** (2017). Tolerating defiance? local average treatment effects without monotonicity. *Quantitative Economics*, 8(2): 367–396

**Duflo, E.**, **Dupas, P.**, and **Kremer, M.** (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in kenya. *American economic review*, 101(5): 1739–74

**Foster, G.** (2006). It's not your peers, and it's not your friends: Some progress toward understanding the educational peer effect mechanism. *Journal of public Economics*, 90(8-9): 1455–1475

**Gottfried, M.A.** and **Harven, A.** (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

**Harris, C.** (2022). Interpreting instrumental variable estimands with unobserved treatment heterogeneity: The effects of college education. *Working Paper*

**Heckman, J.J.** and **Pinto, R.** (2018). Unordered monotonicity. *Econometrica*, 86(1): 1–35

**Hill, A.J.** (2017). The positive influence of female college students on their male peers. *Labour Economics*, 44: 151–160

**Hirano, K.** and **Imbens, G.W.** (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 **Angrist, J.D.** (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2): 467–475

**Jones, T.R.** and **Kofoed, M.S.** (2020). Do peers influence occupational preferences? evidence from randomly-assigned peer groups at west point. *Journal of Public Economics*, 184: 104154

**Lavy, V.** and **Schlosser, A.** (2011). Mechanisms and impacts of gender peer effects at school. *American Economic Journal: Applied Economics*, 3(2): 1–33

**Lucas, R.** (1976). Econometric policy evaluation: A critique. In *Theory, Policy, Institutions: Papers from the Carnegie-Rochester Conferences on Public Policy*, volume 1, 257. North Holland

**Marmaros, D.** and **Sacerdote, B.** (2006). How do friendships form? *The Quarterly Journal of Economics*, 121(1): 79–119

**Rosenbaum, P.R.** and **Rubin, D.B.** (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1): 41–55

**Rubin, D.B.** (1978). Bayesian inference for causal effects: The role of randomization. *The Annals of statistics*, 34–58

**Rubin, D.B.** (1980). Randomization analysis of experimental data: The fisher randomization test comment. *Journal of the American statistical association*, 75(371): 591–593

**Rubin, D.B.** (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

**Sacerdote, B.** (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? In *Handbook of the Economics of Education*, volume 3, 249. Elsevier

**Stinebrickner, R.** and **Stinebrickner, T.R.** (2006). What can be learned about peer effects using college roommates? evidence from new survey data and students from disadvantaged backgrounds. *Journal of public Economics*, 90(8-9): 1435–1454

**Zimmerman, D.J.** (2003). Peer effects in academic outcomes: Evidence from a natural experiment. *Review of Economics and statistics*, 85(1): 9–23

# 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. Among 48 coefficients of interest, four are significant at the 95% confidence level and eight are significant at the 90% confidence level. At both levels of significance, we observe more statistically significant effects than would be expected with perfect randomization of tiebreakers (2.5 and 5). This is cause for concern. However, we note that the baseline characteristics that predict tiebreaker values do not have consistent signs, which suggests a lack of systematic bias

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

|  | 2016 | 2017 | 2018 | 2019 |
|---|---|---|---|---|
|  | (1) | (2) | (3) | (4) |
| Female | 141.7* | -5.682 | -59.01 | 22.71 |
|  | (79.01) | (80.35) | (77.17) | (76.90) |
| STEM | 21.50 | 13.84 | -32.14 | -99.17 |
|  | (80.36) | (80.28) | (77.31) | (75.09) |
| ACT Math | 29.50** | -8.882 | -14.55 | -2.817 |
|  | (12.56) | (12.93) | (11.75) | (11.33) |
| ACT Verbal | -5.628 | 8.257* | 3.163 | -8.251* |
|  | (5.146) | (4.337) | (5.215) | (4.911) |
| Nontraditional | 299.0 | -185.2 | 76.96 | 640.0 |
|  | (312.2) | (275.5) | (319.2) | (489.3) |
| In-State | 127.1 | 149.1* | -71.98 | 127.6 |
|  | (83.82) | (84.39) | (80.23) | (77.76) |
| First Generation | 67.79 | -18.86 | -10.09 | 10.03 |
|  | (105.5) | (106.0) | (91.54) | (100.9) |
| Asian | -373.6* | 282.0 | 375.5** | -342.4** |
|  | (197.0) | (200.6) | (182.9) | (167.5) |
| Black | 185.1 | 204.6 | 439.5 | 381.2 |
|  | (306.8) | (294.6) | (306.1) | (320.6) |
| Hispanic | 88.53 | 97.36 | 270.3 | -307.1 |
|  | (218.0) | (221.9) | (203.8) | (196.8) |
| White | 134.8 | 90.98 | 314.7** | -206.4 |
|  | (150.8) | (152.7) | (137.5) | (135.5) |
| International | 256.4 | 159.5 | 33.19 | -152.2 |
|  | (191.6) | (201.1) | (175.6) | (175.5) |
| Constant | 4119.7*** | 4474.3*** | 4904.9*** | 5515.5*** |
|  | (430.5) | (433.7) | (405.8) | (394.7) |
| Observations | 5688 | 5778 | 5981 | 6269 |

*Notes:* Predictive associations between baseline student characteristics and tiebreakers, by year. *, **, and *** denote significance at the 90, 95, and 99% confidence levels.

# B Room Assignment Monotonicity

As discussed in Section 2.3, 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 mechanical 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 in the complier population who relate to their peers similarly to the way defiers do. 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. We see that 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                        | 0.976            |
| Female                       | 0.976            |
| STEM                         | 0.978            |
| Above Median Math ACT        | 0.977            |
| Above Median Verbal ACT      | 0.977            |
| Nontraditional Student       | 0.958            |
| In-State Student             | 0.979            |
| First Generation Student     | 0.975            |
| Asian                        | 0.983            |
| Black                        | 0.967            |
| Hispanic                     | 0.971            |
| White                        | 0.976            |
| Other Race                   | 0.976            |
| International                | 0.982            |
| No Roommate                  | 0.952            |
| Random Room Assignment       | 0.978            |
| Random Roommate Assignment   | 0.964            |
| Observations                 | 24265            |

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